12-2010

RESISTANCE AND DNA REPAIR MECHANISMS UNDER NITROSATIVE STRESS IN MAMMALIAN AND MICROBIAL SYSTEMS

Hyun-wook Lee
Clemson University, hyunwol@clemson.edu

RESISTANCE AND DNA REPAIR MECHANISMS UNDER NITROSATIVE STRESS IN MAMMALIAN AND MICROBIAL SYSTEMS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Biochemistry & Molecular Biology

---

by
Hyun-Wook Lee
December 2010

Accepted by:
Dr. Weiguo Cao, Committee Chair
Dr. Chin-Fu Chen
Dr. Cheryl Ingram-Smith
Dr. Feng Luo

ABSTRACT

Living organisms are exposed a nitrosative stress mediated by reactive nitrogen species (RNS) that can cause DNA damage and mutation. DNA base deamination is a typical damage occurred under nitrosative stress, which results in conversion of cytosine (C) to uracil (U), adenine (A) to hypoxanthine (I), and guanine (G) to xanthine (X) or oxanine (O). Base excision repair (BER) is an important pathway to remove deaminated DNA lesions in mammalian and microbial systems. My dissertation work concerns with genes and enzymes involved in resistance to nitrosative stress and DNA glycosylases in the BER pathway. In chapter one, I will briefly review current knowledge in these areas. In chapter two, I will present a genetic and biochemical investigation that identifies mouse thioredoxin domain-containing 5 (mTrx 5) and *Escherichia coli* thioredoxin 1 and thioredoxin 2 as genes that are involved in resistance to nitrosative stress. This work indicates radical scavenging as an important resistance mechanism. In chapter three, I will present an extensive biochemical, molecular modeling and molecular dynamics simulations study on deaminated repair activities in *E. coli* mismatch-specific uracil DNA glycoyslase (MUG). Data obtained from cell extracts and purified enzymes indicate that *E. coli* MUG is a robust xanthine DNA glycosylase (XDG) although it is well known as a uracil DNA glycosylase. Site-directed mutagenesis, coupled with molecular modeling and molecular dynamics simulations reveal distinct hydrogen bonding patterns in the active site of *E. coli* MUG, which account for the specificity differences between *E. coli* MUG and human thymine DNA glycosylase, as well as that between the wild type MUG and mutant MUG enzymes. In chapter four, I will describe the deaminated base repair of

DNA glycosylases in archaea. Overall, these studies provide new insights on the cellular mechanisms in resistance to nitrosative stress and deaminated DNA repair mechanisms in mammalian and microbial systems.

DEDICATION

This work is dedicated to my parents, my brother's family and my wife. My parents lead me to the light side all the time since I was born in 1974. Sometimes they were disappointed, angered and saddened by me; however, no matter what may come, they always trust me with all their hearts. My only older brother was my mentor in my youth life and now he is my counselor. Finally, I have found my lifetime companion, that is my wife. She is my strong sole support in life ever and now becomes part of my life. With my full respect, I love you all.

ACKNOWLEDGMENTS

First, I would like to deeply thank my advisor, Dr. Weiguo Cao, who gave me this opportunity to study for academic achievement, supported me and helped me at Clemson University. He has tried to make me improve my scientific insight while no matter what is facing me to discourage myself.

I would also like to thank the members of my graduate committee (Dr. Chin-Fu Chen, Dr. Cheryl Ingram-Smith and Dr. Feng Luo) for their endlessly support and advice throughout my Ph.D. career at Clemson.

Likewise, I would like to thank all the members of the Cao lab (Dr. Thomas Hitchcock, Dr. Honghai Gao, Dr. Rungjuan Mi, Dr. Liang Dong, Bo Xia and Dong-Hoon Lee) for helping in research and giving me more motivation, which made my research easier and all undergraduate students involved in my research projects (Charles Wright, Robert Johnson, Jennifer Kraft) for their assistance.

I would also like to express gratitude to our collaborators, Mr. Allyn R. Brice and Dr. Brian N. Dominy, for their exceptional assistance. I also would like to thank to Dr. Richard P. Cunningham at University at Albany and Dr. Bernard Weiss at Emory University for providing *E. coli* strains.

Finally, I would like to thank my parents again, who have given me unbelievable love, and who have encouraged me not to give up.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

List of Figures (Continued)

Figure Page

ABBREVIATIONS

AAG, Alkyladenine DNA glycosylase

AER, Alternative excision repair

AlkA, 3-methyladenine DNA glycosylase II

AP site, Apurinic/apyrimidinic site

BER, Base excision repair

DSB, DNA double-strand breaks

DNA, Deoxyribonucleic acid

Endo III, Endonuclease III (Nth)

Endo V, Endonuclease V (Nfi)

Endo VIII, Endonuclease VIII (Nei)

Fpg, Formamidopyrimidine-DNA glycosylase

GGR, Global genome repair

H2TH, Helix 2-Turn helix

HhH, Helix-hairpin-helix

iNOS, inducible nitric oxide synthase

MBD4, Methyl-CpG binding protein 4

MED1, Methyl-CpG-binding endonuclease 1

MMR, Methyl-directed mismatch repair

Mpg, Methylpurine glycosylase

M.trx5, Mouse thioredoxin domain-containing 5

MUG, Mismatch-specific uracil DNA glycosylase

Abbreviations (Continued)

Neil, Nei-like

NER, Nucleotide excision repair

NHEJ, Non-homologous end joining

OGG, 8-Oxoguanine DNA glycosylase

RFA, Replication factor A

RNA, Ribonucleic acid

RNI, Reactive nitrogen intermediates

RNS, Reactive nitrogen species

ROS, Reactive oxygen species

RPA, Replication protein A

SMUG1, Single strand-selective monofunctional uracil-DNA glycosylase

SSBs, Single strand breaks

TCR, Transcription-coupled repair

TDG, Thymine DNA glycosylase

UDG, Uracil DNA glycosylase

XDG, Xanthine DNA glycosylase

CHAPTER ONE

DAN DAMAGE AND REPAIR OVERVIEW

I. Introduction

DNA in an organism is subject to alteration constantly by various endogenous and environment factors. The maintenance of an organism's genome integrity is very important for cells in all species. However, Failure of keeping genome integrity will cause DNA mutations, blockage of DNA transactions such as transcription, replication. Therefore, most of organisms including human have evolved a response to prevent deleterious effects of DNA damage, for example, DNA repair mechanisms. Depending on the types of DNA damage, several anti-DNA damage strategies have involved in quenching, removing and eventually repairing those are altered. This chapter provides an overview of thioredoxin's role in reducing DNA damage generated by ROS and/or RNS and followed by relevant DNA repair pathways on both types of typical DNA mutations. Also, it will address the relationship of DNA repair deficiencies to diseases in human.

II. Anti-ROS and Anti-RNS mechanism of thioredoxin

Cells are always confronted to multiple factors that are able to damage DNA, including Reactive oxygen species (ROS) and Reactive nitrogen species (RNS). Both species have emerged as ubiquitous signalling molecules participating in the recognition of and the response to stress factors and are regarded as damaging agents to cells as well (1). In order to detoxifying ROS/RNS, Cells are able to reduce damage using several enzymes such as superoxide dismutases, glutathione peroxidases and thioredoxin.

A. Reactive oxygen species (ROS)

Reactive oxygen species (ROS) are molecules that are generated continuously in most cells. ROS can interact with various macromolecules and thus cause damage to DNA (2). In this way, the major ROS (superoxide anion, hydrogen peroxide, and hydroxyl) are formed by the sequential reduction of O_2 (3, 10) (Table 1-1). In aerobic organisms, products of many catabolic pathways are ended up with O_2 in the mitochondria in order to produce energy but also including ROS as by-products at the mitochondria (3). ROS are generated by partial reduction of O_2, which generates superoxide ($O_2^{\bullet -}$) or hydrogen peroxide (H_2O_2) and can also be generated if O_2 interacts with upstream complexes, particularly complexes I and III of electron transfer chain in mitochondria (3, 4). Several enzymatic systems in the cytosol also generate H_2O_2, for example, amino acid oxidases, cyclo-oxygenase, lipid oxygenase, xanthine oxidase and notable sources of ROS at the plasma membrane are the NADPH oxidases (3).

B. Reactive nitrogen species (RNS)

Nitric Oxide ($NO^{\bullet}$) plays a role in the nitrosative stress as a main RNS produced by cells and/or as a main source for the other RNS. $NO^{\bullet}$ is produced by NO synthases (NOS) from L-arginine and oxygen (7). There are three isoforms of NOS. NOS 1 or nNOS (for neuronal), NOS 2 or iNOS (for inducible), and NOS 3 or eNOS (for endothelial) (8). Both nNOS and eNOS are constitutively expressed when iNOS expression is induced, predominantly in macrophages, under pro-inflammatory conditions, large quantities of $NO^{\bullet}$ are produced (9). Along with ROS, RNS and in particular nitric oxide ($NO^{\bullet}$), are now considered as major components of redox state

regulation although these are limited to consider as damaging agents. Indeed, $NO^{\bullet}$ can participate in specific signal transduction pathways which has been shown to be a

Table 1-1 The Major ROS Molecules and their metabolism. Adopted from (10).

ROS molecule	Main sources	Enzymatic defense systems	Product(s)
Superoxide ($O_2^{\bullet -}$)	'Leakage' of electrons from the electron transport chain Activated phagocytes Xanthine oxidase Flavoenzymes	Superoxide dismutase (SOD) Superoxide reductase (in some bacteria)	$H_2O_2 + O_2$ H_2O_2
Hydrogen peroxide (H_2O_2)	From $O_2^{\bullet -}$ via superoxide dismutase (SOD) NADPH-oxidase (neutrophils) Glucose oxidase Xanthine oxidase	Glutathione peroxidase Catalases Peroxiredoxins (Prx)	$H_2O + GSSG$ $H_2O + O_2$ H_2O
Hydroxyl radical ($^{\bullet}OH$)	From $O_2^{\bullet -}$ and H_2O_2 via transition metals (Fe or Cu)		
Nitric oxide (NO)	Nitric oxide synthases	Glutathione/TrxR	GSNO

ubiquitous signalling molecule in mammalian and plants (1,5). $NO^{\bullet}$ can interact with other radical ions such as $O_2^{\bullet -}$,and thus become a peroxynitrite (ONOO-). RNS refers to various nitrogenous products, such as $NO^{\bullet}$, nitroxyl (HNO), nitrosonium cation (NO^{+}), higher oxides of nitrogen, S-nitrosothiols (RSNOs), $ONOO^{\bullet}$, and dinitrosyl iron complexes (6, 7) (Table 1-2).

C. Anti-ROS/RNS mechanism of thioredoxin.

Thioredoxin 1(TrxA), a small ubiquitous 12kDa protein with a conserved active site sequence (Cys-Gly-Pro-Cys), was initially identified in *Escherichia coli* as an electron donor for ribonucleotide reductase (11). Thioredoxin is part of thioredoixn systems which include NADPH and thiroredoxin reductase. Recently, the second thioredoxin, *E.coli* Trx2 (TrxC) was identified and have additional N-terminal domain of 32 amino acids including two CXXC motifs. Human thioredoxin1 (TRX1) was cloned as an adult T cell leukemia (ATL)-derived factor produced by HTLV-I transformed T cell

line ATL-2 cells (12, 13). Human thioredoxin 2 (TRX2) also was identified in mitochondria (14). Thioredoxins are highly conserved from bacteria to mammalian.

Table 1-2 Main reactive nitrogen species. Adopted from (7).

Name	*Formula*	*Formation*
Dinitrogen trioxide	N_2O_3	From NO· and O_2
Nitric oxide or nitrogen monoxide	NO·	From NOS
Nitrite	NO_2^-	From NO·
Nitrogen dioxide	NO_2·	From $ONOO^-$decomposition
Nitronium cation	NO_2^+	From $ONOOCO_2^-$decomposition
Nitrosonium cation	NO^+	From NO·
Nitrosoperoxycarbonate anion	$ONOOCO_2^-$	$ONOO^-$ and CO_2
Nitroxyl	HNO	From one-electron reduction of NO·
Nitryl chloride	Cl-NO_2	From NO_2^- and HOCl
Peroxynitrite	$ONOO^-$	From NO· and $O_2\cdot^-$
S-Nitrosothiols	RSNOs	From covalent addition of an NO· group to a cysteine thiol/sulfhydryl

Thioredoxins from both prokaryotic and mammalian organisms are reduced by mammalian TrxR, while *E.coli* TrxR has a narrow substrate specificity and only catalyzes the reduction of bacterial Trx (10, 11, 18, 19, 20, 21). Thioredoxin has specific functions for protecting cells from toxicant, especially reactive species (10, 22) (Table 1-3). Besides thioredoxins that exist in all species, proteins that share the similar active site sequence as TRX (Cys-X–Y–Cys) are also identified in mammalian systems (15, 16, 17). Given that ROS/RNS play a role as a second messanger or as an oxidative/nitrosative stress causing agent, we have to consider their detoxification. ROS/RNS detoxification relies on the antioxidant defence (AD), which involves enzymatic activities (catalases, superoxide dismutases, peroxidases) and antioxidant molecules (glutathione, GSH, ascorbate, thioredoxin) (1). Taken together, ROS/RNS and AD contribute to the redox balance that regulate many protein's activity such as transcription factor (10) and that reduce DNA damage possibly such as DNA oxidation and DNA deamination.

Table1-3 Functions of thioredoxin-1. Adopted from (22)

Cell proliferation; growth factor
DNA synthesis
Ribonucleotide reductase
Redox signaling
Regulation of transcription factors (NF-kB, AP-1 [Ref-1], p53)
Detoxification; antioxidant
Reduction of H_2O_2; thioredoxin peroxidase, peroxiredoxins
Protein reduction
Methionine sulfoxide reduction
Protein disulfide reduction
Apoptosis
Reduced form of Trx-1 complexes with apoptosis signal-regulating kinase (ASK1); this prevents downstream signaling
Trx-1 likely affects other apoptosis pathways

Among various functions of thioredoxin, I will focus on reduction/oxidation reaction (redox) signaling for proteins and reactive species, in particular, the specific mechanism through which Trx may be exerting its modulating effects. Thioredoxins 1 and 2 play a direct role in reducing oxidative species, as have been illustrated for hydrogen peroxide (23, 24). Therefore, the ability of Trx-1 to reduce thiols is critical for DNA binding on nuclear transcription factors such as AP-1 and NF-κB (29) and provides protection against oxidative stress, bleomycin-induced lung damage, and doxorubicin-induced cardiotoxicity (22, 26, 27, 28). Thioredoxin's 2 cysteines at the active site that are highly conserved in the sequence (Cys-Gly-Pro-Cys) are oxidized by forming a disulfide bond between two cysteines while the target protein is reduced by making disulfide bond between transferring reducing equivalents of trx and cysteine of target protein (25). The oxidized active site of Trx-1 and Trx-2 are reduced by TR1 (thioredoxin reductase 1) and TR2 (thioredoxin reductase 2), respectively, using electrons from NADPH. According to the effect of trx on oxidative stress, we might also consider whether trx will influence nitrosative stress mediated through $NO^{\bullet}$, nitroxyl (HNO),

nitrosonium cation (NO^+), higher oxides of nitrogen, S-nitrosothiols (RSNOs), $ONOO^{\bullet}$, and dinitrosyl iron complexes (6, 7) (Table 1-2). Signaling from Nitric Oxide ($NO^{\bullet}$) can generate S-nitrosylation of proteins including trx. Many redox-related proteins have active cysteines which are able to S-nitrosylation such as catalase (35), glutathione peroxidase (36), glutathione reductase (37), glutathione transferase P1–1 (38), and thioredoxin (34). The S-nitrosylation reaction can be mediated by a direct reaction between NO and thiols in the presence of electron acceptors (30). The nitrosation of sulfhydryl (SH) groups via acidified nitrite or NO/O2 has been studied (31, 32). According to previous studies, researchers have showed that thioredoxin-1 can transfer NO to caspase-3 and thereby inactivating the enzyme (33, 34). They also found that purified thioredoxin-1 is S-nitrosylated at cysteine 73 and not at cysteine 69 (33) (Fig. 1-1), which is different from the claim of Dimmeler's group that thioredoxin-1 is S-nitrosylated at cysteine 69 (34). Given the fact that thiroedoxin-1 can be S-nitrosylated, both, cysteine 69 and 73 residues, are potential site for S-nitrosylation. On the contrary, it has been found that S-nitrosylation of thioredoxin-1 at active-site Cys32/Cys35 leads to the dissociation and activation of apoptosis signal-regulating kinase 1 (ASK1) (39).

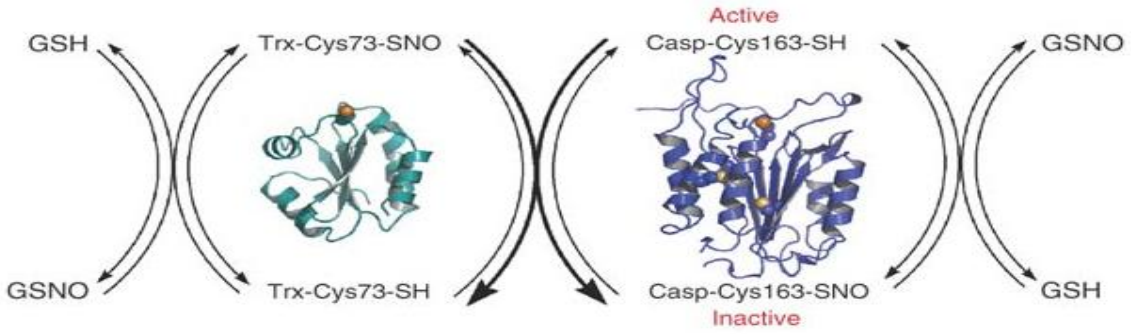

Figure 1.1 Schematic representation of S-Nitrosylation of C73 of Thioredoxin by excess GSNO. Adopted from *http://www.cchem.berkeley.edu/mmargrp/research/SNO/SNO.html*

Taken together, thioredoxin-1 may play a role in endogenous protective mechanism for cells against oxidative/nitrosative stress and regulate the target proteins by means of S-nitrosylation.

III. DNA Damage on oxidative/nitrosative stress

DNA damage plays a role in mutagenesis and carcinogenesis. Cells undergo constant and extensive damage from endogenously through hydrolysis, exposure to reactive species. It causes DNA modifications, in particular DNA oxidation and DNA deamination. It is important to understand that what kind of DNAs are being mutated, how cell could prevent DNA mutations from endogenously generated events and finally what happen if cell cannot take care of DNA damage. This part of the chapter will present oxidatively modified DNA and nitrosatively modified DNA.

1. Damage to DNA by reactive oxygen species (ROS)

DNA base modifications are generated by both exogenous and endogenous factors. Especially, Endogenous DNA damage occurs at a high rate compared with exogenous damage. The types of damage produced by endogenously processes are the same as by exogenous (40).

1. Sources of DNA oxidation

Oxidation of DNA can be generated by a variety of factors such as UV light, ionizing radiation, cigarette smoke including endogenous cell metabolism. This process may include hydroxyl radical, singlet oxygen, hydrogen peroxide and one-electron

oxidation. They are made as by-products of the respiratory electron transport chain, by cytochrome P450 and xanthine oxidase metabolism, by micorsomes and peroxisomes, and are also produced by neutrophils, eosinophils and macrophages during inflammation and in various metal-catalyzed reactions (41, 42).

2. Types of DNA oxidation

Oxidation is the major contributor to DNA base damage, with estimates of the total number of oxidized bases formed approximating 10,000 adducts per cell per day among several different lesions such as base- or deoxyribose lesions, strand breaks and cross-links, base modification. (41, 43).

a. 8-oxo-7,8-dihydroguanine (8-oxoG)

Hydroxyl radicals can react to guanine and adenine at positions 4, 5, or 8 in the purine ring (41.44). 8-oxo-7,8-dihydroguanine (8-oxoG), also called 8-hydroxyguanine, is regarded as the most abundant oxidative damage and a marker of cellular oxidative stress. The addition of hydroxyl radical to C-8 of guanine produces a C-8 OH-adduct radical, which can also be oxidized to 8-oxoG (41, 44) (Fig. 1-2). Two or three residues of 8-oxoG are present per 10^6 G sites in human leukocytes and roughly 80 8-oxoG residues are continuously generated per human cell per day (41, 45, 46).

b. 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG)

The formation of FapyG is also generated by reducing C-8 OH-adduct radical in guanine which leads to imidazole ring-opening (41, 44).

c. 5-hydroxycytosine (5-OH-C) and 5-hydroxyuracil (5-OH-U)

Cytosine is oxidized only at the 5,6-double bond and main oxidative cytosine modifications found in DNA are 5,6-dihydroxy-5,6-dihydrocytosine (cytosine glycol, Cg), its deamination and dehydration products 5,6-dihydroxy-5,6-dihydrouracil (uracil glycol, Ug) and 5-hydroxycytosine (5-OH-C), 5-hydroxyuracil (5-OH-U) (formed from Ug by dehydratation or from 5-OH-C by deamination) (44, 47, 48).

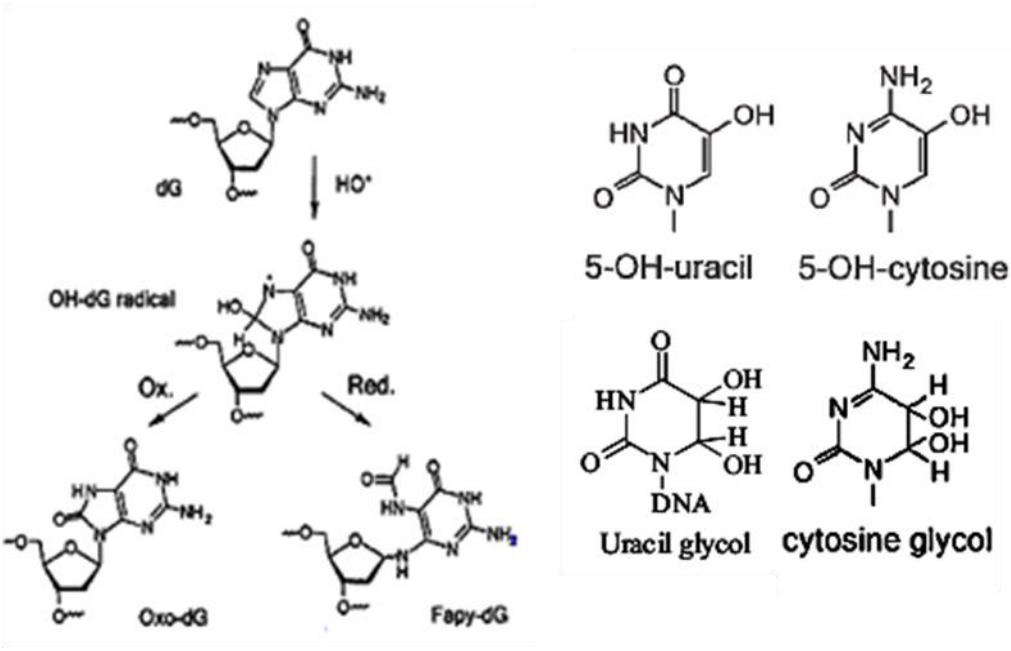

Figure 1.2 Modified bases resulting from the attack of reactive oxygen species. Adopted from (41, 44).

2. Damage to DNA by reactive Nitrogen species (RNS)

Exposure of DNA to reactive nitrogen species can promote deamination of DNA bases by changing of guanine to xanthine (X) and oxanine (O), adenine to hypoxanthine (I), cytosine to uracil (U), 5-methylcytosine (5-meC) to thymine. Spontaneous deamination is rather low, however, it can be increased in vivo by nitrogen dioxide (NO_2) formed during inflammation or by UV (49).

a. Deamination of cytosine

The extracyclic amino group at position C4 of cytosine is unstable, and is slowly lost by hydrolysis under physiological conditions to form uracil (55). It can occur spontaneously or as a result of treatment with acidified sodium nitrite. This reaction results in the deamination of cytosine to uracil which can be then base paired with adenine. Due to the fact that uracil is a normally occurs in RNA, this causes potentially a problem in cells. The deamination of cytosine leads to G=C→A= T point mutation and G→A and C→T transitions (50, 51, 52, 53, 54) (Fig. 1-3, 1-4). Most uracils in DNA are recognized by uracil-DNA glycosylase, but the AP-site generated is potentially mutagenic and cytotoxic unless repaired (54). Some studies estimate cytosine deamination at ~70–200 events per human cell per day (49). The rate of deamination occurs 200–300 fold faster from ssDNA than from double-stranded DNA (dsDNA) (45).

5-Methylcytosine can be deaminated to form thymine with use of reagents such as nitrous acid which can be also G→A and C→T transitions (54) (Fig. 1-3, 1-4). It is found in the DNA of all studied vertebrates and higher plants (57, 58) and mostly,

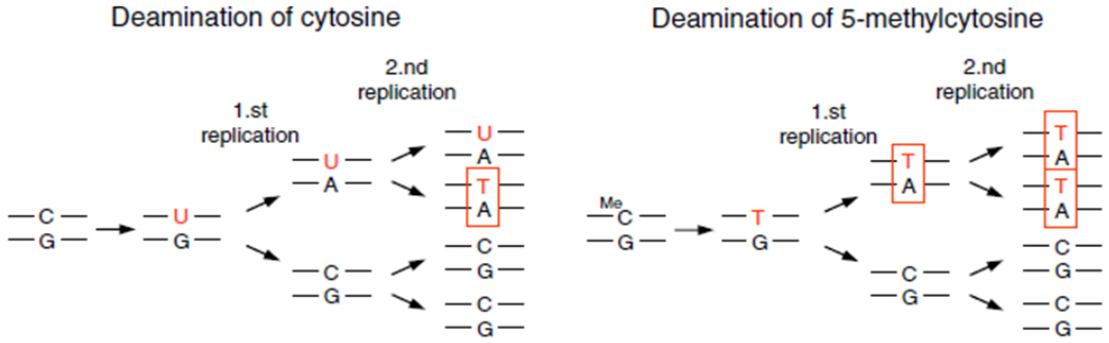

Figure 1.3 The subsequent base pairing following the entry of uracil and thymine. Adopted from (54).

DNA methylation in higher plants can be found more than in mammals (up to approximately one-third DNA is to help control gene expression (59, 60).

b. Deamination of adenine

Deamination of adenine can occur at natural condition, except at much lower frequency than that of cytosine deamination (64). The deamination of adenine leads to the reverse A→G and T→C transitions (61, 62, 53) (Fig. 1-5, 1-6). In vivo, purine nucleotides are degraded by a pathway such as adenosine deaminase which can deaminate adenosine to inosine. Since inosine is naturally and specifically generate within the cell,

Figure 1.4 Schematic representation of the deamination of cytosine and 5-methyl cytosine.

it can also be integrated into the chromosomal DNA under specific environment such as rdgB deletion mutant (65). Inosine is allowed to incorporated to genomic DNA or adenosine is accumulated by lack of adenosine deaminase, those phenomena are indeed detrimental.

c. Deamination of Guanine

The mechanisms of nucleobase deamination result in the formation of xanthine from guanine. In addition, the reaction of nitrous acid with guanine in vitro has been

shown to partition to form both xanthine and oxanine as well (66). The favor for

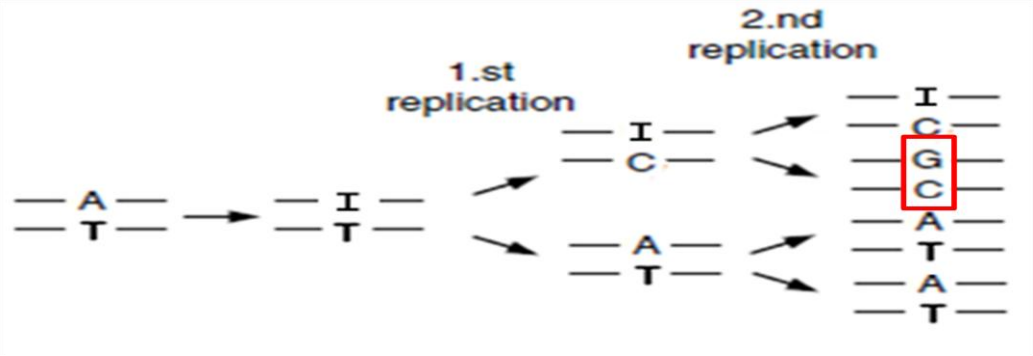

Figure 1.5 The subsequent base pairing following the entry of inosine.

hydrolytic deamination occurs in the order 5-methylcytosine > cytosine > adenine > guanine (67, 68).

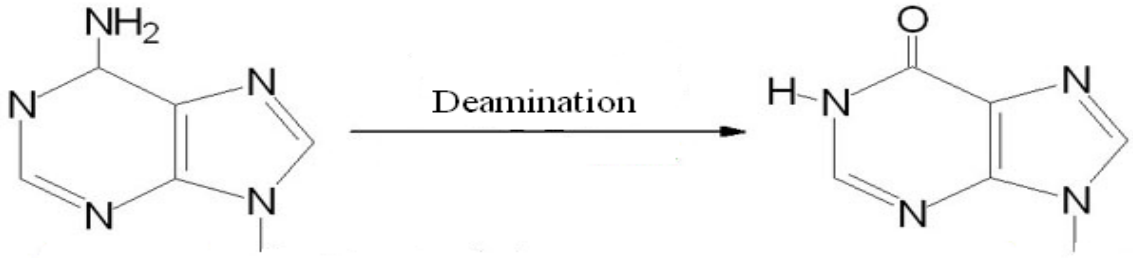

Figure 1.6 Schematic representation of the deamination of adenine.

The reactivity of nitrous anhydride (N_2O_3) with xanthine formation proposed to occur at twice the rate of uracil (69) and hypoxanthine (70). In human lymphoblastoid TK6 cells interacted with NO, the levels of both xanthine and hypoxanthine accerated ~40-fold over untreated (70). Deamination of xanthine lead to G:C→A:T mutations (71, 72) (Fig. 1-7, Fig. 1-8). Recently other groups presented evidence that under biological conditions dX is a relatively stable lesion with a half-life of 2 years at 37°C and pH 7 in double-stranded DNA (73, 74), although xanthine in DNA under acidic conditions is unstable due to depurination (66, 75).

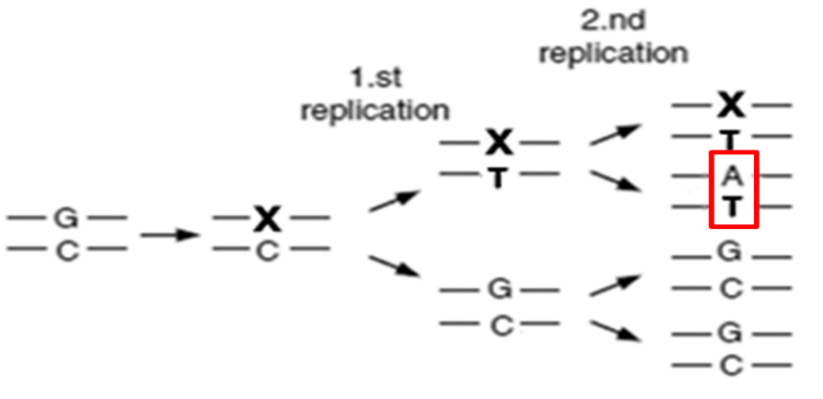

Figure 1.7 The subsequent base pairing following the entry of xanthine.

Given the steady-state kinetics of dNTP incorporation opposite X, preferred pairing for X is the order of T > C >> A ≈ G with *Drosophila* DNA polymerase α (71). The preferences for HIV-1RT were C ≈ T > G > A, whereas for DNA PolI(KF$^-$) is C > T, with no discernible incorporation of either dATP or dGTP (77). Oxanine (5-amino-3-β-D-ribofuranosyl-3*H*-imidazo[4,5-*d*]oxazin-7-one) was originally isolated from *Streptomyces capreolus* MG265-CF3 and has a antibacterial activity (81).

In the early 1980s, it was also shown that oxanine inhibits the growth of leukemia cell lines in addition to nucleic acid synthesis and *in vitro* cell growth (82).

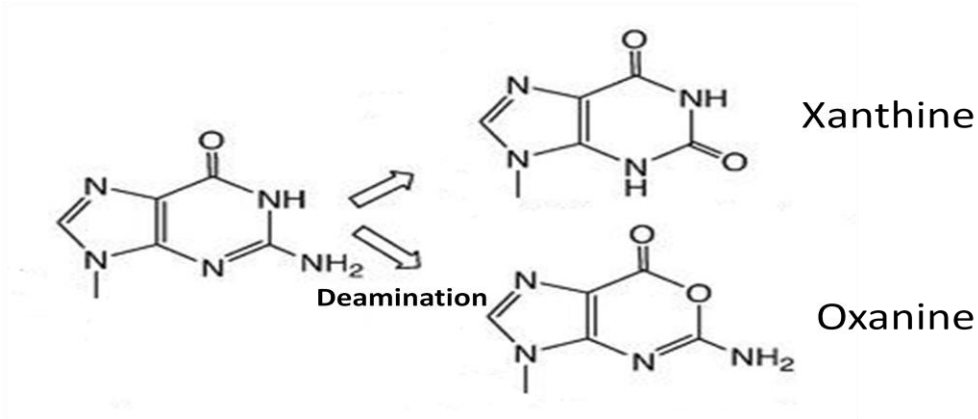

Figure 1.8 Schematic representation of the deamination of guanine.

Oxanine is, together with xanthine, a major product of guanine deamination when guanine reacts with N_2O_3 with the molar ratio of 1:3. It has been studied that the N-glycosidic bond of oxanine is as stable as that of guanine and hydrolyzed 44-fold more slowly than that of xanthine (66). DNA polymerase can incorporate cytosine and thymine into opposite oxanine strand (78) and thus deamination of oxanine lead to G:C→A:T mutations (83) (Fig. 1-8, 1-9), but if oxanine exist in the template DNA, it is able to pair with four bases, which means oxanine in DNA is more mutagenic (80).

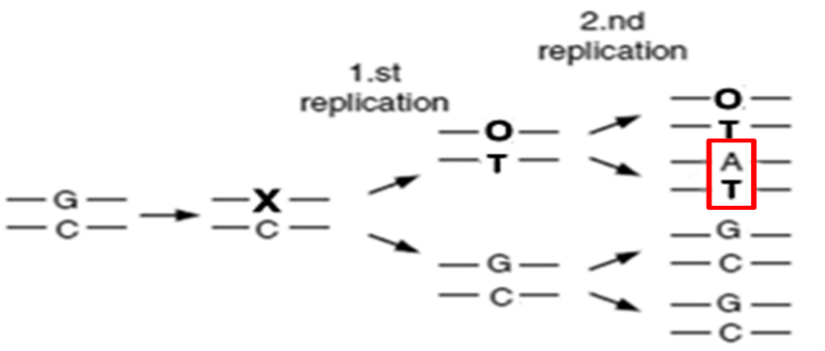

Figure 1.9 The subsequent base pairing following the entry of oxanine.

As oxanine react efficiently with animo groups, it is easily to form cross-link adducts with amino acids (79). It could also interact with peptides, proteins and other biomolecules.

IV. DNA repair mechanisms

Cellular DNA is subjected to various modifications resulting from endogenous and environmental sources such as oxidation, deamination, depurination, DNA adducts, as well as mismatch base pairs. Due to the existence of a variety of DNA damage, it requires several DNA repair mechanisms to protect cells from harmful consequences. It

depends on the types of damage, some organisms have evolved specific repair pathways or evolved broad range of repair machineries. Here we will discuss each of the types of DNA repair pathways, in addition to proteins involved in each respective pathway.

A. Base excision repair pathway

Once the damage has been detected, specific repairing molecules are recruited to on or near the damaged site. If damage is small or simple, cell is using BER (Base excision repair) system. Base excision repair (BER) is one of the major active DNA repair processes that allows the efficient recognition of damage site and removal of a damaged DNA base. Each of these lesions is either cytotoxic or mutagenic so that these are repaired via the BER pathway in organisms ranging from Escherichia coli to mammals (84, 85, 86, 87). In this pathway, damaged bases are removed by enzymes called DNA N-glycosylases. This is done by hydrolysis of its N-glycosidic bond that links the damaged base to deoxyribose-phosphate backbone of the DNA. The resulting apurinic/apyrimidinic (AP) sites are recognized by AP endonuclease (for monofunctional glycosylase) or N-glycosylases obtaining an AP-lyase activity (bifunctional enzyme). The gap which may require an editing reaction to generate unblocked 5′ and 3′ ends, are then filled by DNA polymerase, either with a single-nucleotide, short patch or with a longer repair patch, followed by a ligation step (44).

In mammalian cells, short-patch BER (SP-BER) is the repair that involves the replacement of a single nucleotide mediated by POL β, which incorporates the correct nucleotide and removes the 5′-dRP terminus through its dRPase activity and followed by LIG3/XRCC1 complex for ligation step (44) (Fig. 1-10).

Sometimes it needed to incorporate beyond a single nucleotide, more than two to 10 nucleotides. This alternative pathway is known as long-patch BER (LP-BER) and being used less often. The longer-patch pathway depends on enzymes normally involved in DNA replication: DNA polymerase δ (Pol δ) or ε (Pol ε), replication factor C (RFC), and proliferating cell nuclear antigen (PCNA), which act as cofactors for both POL δ and POL ε (86, 88). The replaced strand is displaced and cut away by DNAase IV or flap endonuclease 1 (FEN1) and followed by the ligation step that is performed by LIG1 (44, 86) (Fig. 1-10).

A. Repair of oxidation

1. Oxidized purine glycosylases (FPG/Ogg1)

Fpg belongs to the H2TH superfamily and while Ogg1 belongs to HhH DNA glycosylases. Ogg1 was initially discovered in *S. cerevisiae* by functional complementation (96, 100) and is also found in mammals, fungi and unicellular eukaryotes (101).

Human Ogg1 is a bifunctional enzyme and has narrow range of specificity such as 8-oxoG pair with C, faPyG, me-faPyG and 7,8-dihydro-8-oxoadenine (100, 102). In archaea, hyperthermophilic *Pyrobaculum aerophilum* has member of a new family of HhH DNA glycosylases, Agog which is able to excise 8-oxoG in single-stranded DNA (ssDNA) as efficiently as in double-stranded DNA (dsDNA) with no preference (103).

Escherichia coli Fpg (also called MutM) excise 8-oxoG mismatched with cytosine, faPyG and bifunctional enzyme. Fpg has no sequence homolog in eukaryotic, but is functionally similar to Ogg1 homolog.

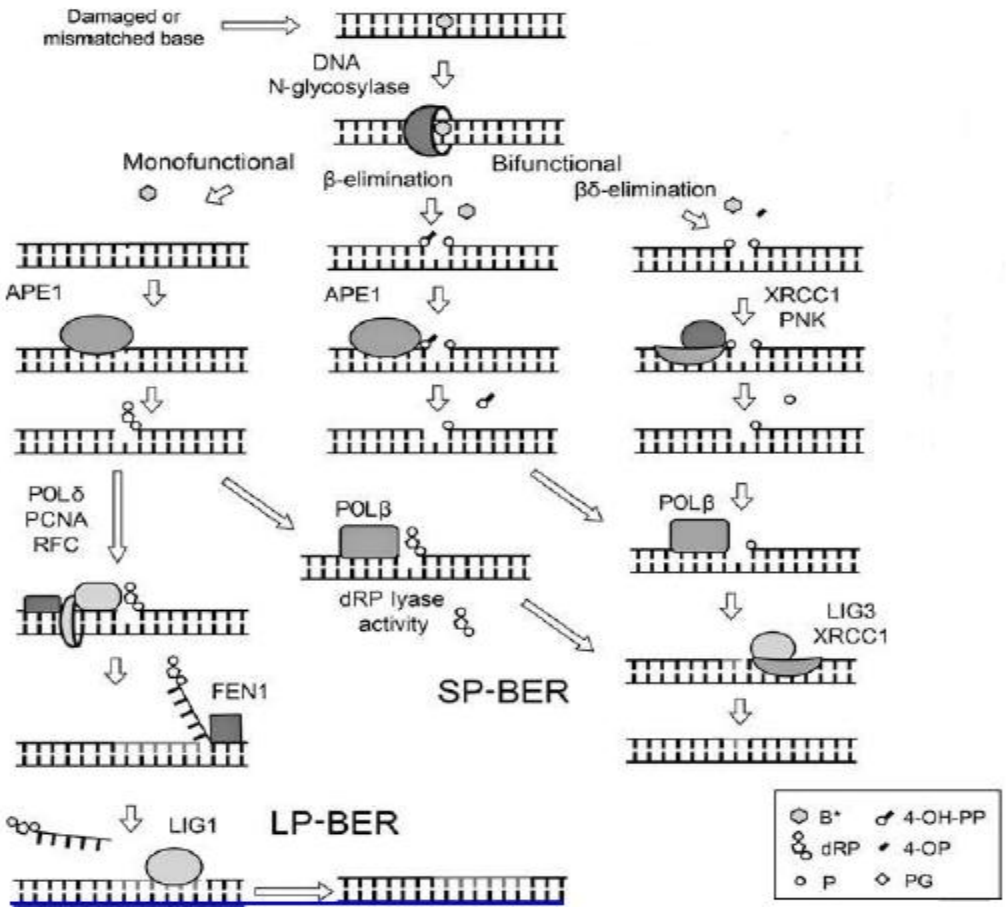

Figure 1.10 Schematic representation of the BER pathway. adopted from (44).

2. Oxidized pyrimidine glycosylases (Nth/Nei)

Nei belongs to the H2TH superfamily, while Nth belongs to HhH DNA glycosylases and both are bifunctional DNA glycosylases.

Endonuclease III, is encoded by the *nth* gene, first identified from *Escherichia coli*, originally referred as a DNA nicking activity seen after heavy ultraviolet irradiation (89). From bacteria to human, Enzymes with similar substrate specificities have been found in other bacteria (90), in yeast (91) and in mammalian cells including human (92, 93, 94). *E.coli* MutY recognizes adenine from a pair with 8-oxoG shows significant

homology to endonuclease III (95) and monofunctional enzyme. This 23.5 kDa protein recognizes a large number of oxidized pyrimidines such as thymine glycol, urea, 5,6-dihydroxythymine, 5,6-dihydrouracil, 5-hydroxy-5-methylhydanton, uracil glycol and 5-formyluracil (96).

Escherichia coli Endonuclease VIII (Nei) is a paralog of Fpg with capacity of β,δ-elimination. The enzyme also has broad substrate specificity such as thymine glycol, 5-hydroxy-5-methylhydanton, 5,6-dihydroxythymine and urea (97, 98). In human, a family of three proteins (Endonuclease VIII like proteins, NEIL1, NEIL2 and NEIL3) that showed homology to the *Escherichia coli* Fpg/Nei DNA glycosylases was identified.

B. Repair of deamination

1. UDG superfamily

Uracil is a normal component at the cellular level and is component of RNA, however, uracil itself is able to incorporate into DNA by cytosine deamination and eventually, it cause mutagenic effect if not repaired.

Initially, the enzyme for the repair of uracil was discovered 1974, which was purified from *Escherichia coli* (104). This is now called uracil-DNA glycosylase (UDG or UNG) and belongs Family 1 in the UDG superfamily. Here, we will discuss 5 members of UDG superfamily.

a. Uracil DNA glycosylase (UNG) Family 1

UNGs recognize base uracil from single-stranded DNA and from double-stranded DNA which is highly specific for uracil in DNA. Apparently, it has base preference in the order of SSU=U/A>U/G>>U/T or U/C or uracil in RNA. UNGs typically exist among

prokaryotes and eukaryotes and even in some DNA viruses, however, it is absent to archaea. In humans, two UDG activities are found which were located to either nuclei or mitochondria through alternative splicing of the same gene (105). UNG1 is in mitochondria (105) and UNG2 is in nuclei which can interact with RFA (replication factor A) and proliferating cell nuclear antigen (106). N-terminal regions can determine the localization, but the catalytic domain is the same for both nuclear and mitochondrial forms. UNG2 play a role in rapid removal of dUMP residues during DNA replication and regulate post-replicative removal of misincorporated uracil in DNA. The *ung*-/- cells show a strong deficiency in removal of mis-incorporated uracils and have a steady-state level of approximately 2000 uracil residues per cell (107). UNG mRNA levels increase 8–12-fold late in G1 phase while activity reaches a maximum in early S phase (108).

Crystal structures of viral (109), human (110), and more recently *E. coli* (111) UNGs show that it has a highly conserved that are five major structural motifs first, the minor-groove intercalation loop; second, the Pro-Rich loop; third, the Gly-Ser loop; fourth, the uracil-specificity region; and last, the water-activating loop. These motifs play a critical role in detection of uracils in DNA (111, 112).

b. Uracil DNA glycosylase (UDG) Family 2

Thymine DNA-glycosylase (TDG) and mismatch-specific uracil-DNA glycosylase (MUG) belong to Uracil DNA glycosylase (UDG) Family 2 that removes uracil and thymine from G:T for TDG (113, 114) and are highly specific on G:U mismatches in double-stranded DNA (dsDNA) for MUG (115). Both enzymes are insensitive by Ugi which can inhibit Family-1 UDG. The crystal structure of the *E. coli*

MUG enzyme indicated that a similar structural homology to the Family-1 UDGs although they have only 10% sequence identity (116). In particularly, two highly conserved motifs, the water activating loop (GQDPY) and the minor groove-intercalating loop (HPSPLS) in UNG, are equivalent to in MUG corresponding to GINPGL (identical in human TDG) and NPSGLS (MPSSSS in human TDG) (88). The asparagine (underline) in MUG cannot activate the nucleophilic water, however it is able to bind and present a water molecule in almost exactly the same position in the active site, as the aspartate in UNG (88). TDG may participate in the demethylation of CpG sites by processing G·T mispairs created by deamination of 5-methylcytosine which would dramatically increase G·T mispairs (117, 118). Most glycosylases use a bulky side chain (Asn, Arg, Gln, Leu, or Tyr) to intrude the helical space (empty space) created by nucleotide flipping, except 3-methyladenine DNA glycosylase which has a Gly at this position. According to the crystal structure of TDG bound to abasic DNA (119) , TDG uses Arg^{275} side chain to penetrate the minor groove, intruding the space vacated by the flipped abasic nucleotide and contacting two DNA backbone phosphates that flank the target site. Interestingly, Drohat group reveals that the crystal structure of hTDG prefers interaction for guanine than adenine as the pairing partner of the target base and interactions that likely confer CpG sequence specificity (119, 120). The N-terminus of TDG can interact with DNA in a non-sequence-specific manner to process G:T mispairs while the catalytic domain scans for G·U mismatches, resulting in a slow enzymatic turnover. Moreover, the N-terminus performs regulatory functions and contains the sites for interaction and posttranslational modification by transcription-related activities while the C-terminus region is modified

by covalent conjugation of small ubiquitin-like modifier (SUMO) protein (121, 124, 125). According to a recent study, hTDG might be also involved in transcription of active regions of the genome and thus BER may be coupled to transcription (122). Interestingly, overexpression of TDG has shown to re-activate a hormone-regulated transgene silenced by CpG methylation, suggesting a role for TDG in epigenetic regulation (123).

c. Uracil DNA glycosylase (UDG) Family 3

SMUG1 was initially found in *Xenopus laevis*, as a single-strand selective monofunctional uracil-DNA glycosylase. More characterization of SMUG1 had been determined in human SMUG1. Human SMUG1 also can removes uracil efficiently from both U:G mismatches and U:A base pairs in addition to be active on 5-hydroxymethyluracil, 5-hydroxyuracil, and 5-formyluracil (126, 127, 128), Recently, some groups have extended the substrate range of hSMUG1 to include CaU which is thymine methyl group oxidation (129). These results explain the hSMUG1 could play a role in the repair of DNA oxidation damage in vivo. Analysis of SMUG1 sequences revealed some similarity with those of the Family-1 and Family-2 UDG, in particularly in the motifs associated with their active sites, however, it should also be noted that SMUGs are not recognised as structural homologues of UDGs or MUG/TDG. SMUG1 was thought to be present only in vertebrates, however, indeed it is also found in some bacteria such as *Geobacter metallireducens* (Gme), *Azoarcus* species (Asp), *Rhodopirellula baltica* (Rba), and *Opitutaceae* bacterium (Oba) (130). Interestingly, bacteria contain either the UNG-type enzyme or the SMUG1-type enzyme, usually not both. Studies on *ung* knockout, *Smug1* knockdown and *ung* knockout/*Smug1* knockdown

mouse cells have indicated that Smug1 and Ung2 are both required for the prevention of mutations and that their functions are not redundant (131, 132). Although SMUG1 and UNG2 have similar prospective in vivo, both are quite different enzyme such that SMUG1 cannot reduce mutation rates which is induced by AID, unlike UNG2 which alleviates the effects of AID (132). The crystal structure of xSMUG1–DNA complex reveals that it penetrates the double helix with a wedge motif that binds tightly to the abasic site with the strongest binding to AP:G (133). When SMUG1 is over-expressed, due to SMUG1's strong attachment to AP-sites, it may interfere with replication, and thus prevent cell division (134).

d. Uracil DNA glycosylase (UDG) Family 4

Family 4 UDGs, which are specially found in thermophilic bacteria and archaea, possess a 4Fe–4S cluster. Sandigursky and Franklin have discovered uracil-DNA glycosylases encoded in the genomes of the thermophiles *Thermotoga maritima* from bacteria (137) and *Archaeoglobus fulgidus* from archaea (138) which appear to be different from other UDG families. The enzyme activity from both species are capable of removing uracil from double-stranded DNA containing either a U/A or U/G base pair as well as from single-stranded DNA.

Miller and colleagues characterized *Pyrobaculum aerophilum* DNA glycosylase (Pa-UDGa) that removes uracil and thymine from G·U/ G·T and its inhibition by the uracil-DNA glycosylase inhibitor (Ugi). Interestingly, its sequence is closely related to the *Thermotoga maritima* (Tma UDGa) and *Archaeoglobus fulgidus* (Afu UDGa) enzymes (139).

Another thermophilic bacterium, *Thermus thermophilus* UDG (Tth UDGa) processes UDG activity on both single-stranded and double-stranded DNA, regardless of opposing base, but does not have G:T mispair activity (140).

e. Uracil DNA glycosylase (UDG) Family 5

Pyrobaculum aerophilum's genome database was available in 2002, Pa-UDGb was initially found as a family 5 that can remove hypoxanthine (136). Recently, another family of UDGs was found in *Thermus thermophilus HB8*, which is called Tth UDGb and missed a polar residue, corresponding to a catalytic residue, glutamate in the family 4 UDGs. These novel Tth UDGb, designated family 5, have broad substrate specificity such as G/T mismatch, 5'-hydroxymethyluracil, 5'-fluorouracil and a distinctive sequence motif in the active site (135). The iron–sulfur (4Fe–4S) cluster was also found in family 5 UDG as in family 4 UDGs (135).

3. AlkA (3-methyladenine DNA glycosylase II)

The 3-mA DNA glycosylase I (*E. coli* Tag), 3-mA DNA glycosylase II (*E. coli* AlkA) and the *Helicobacter pylori* 3-mA DNA glycosylase III (MagIII) all belong to the HhH superfamily (141, 142). AlkA uses a loop between two α-helixes to wedge into the minor groove of the DNA, while Aag uses a β-hairpin motif (142). Incucible *E. coli* AlkA has a wide range of substrate specificity such as excising bases N3- and N7-substituted purines, O2-substituted pyrimidines, εA and εC, hypoxanthine (142). Hypoxanthine was initially found as a substrate for AlkA that has different substrate specificity. According to the crystal structure of AlkA in complex with DNA duplex containing an abasic site

analog, 1-azaribose, the catalytic site Asp238 is positioned close to the flipped abasic site, and Leu125 is inserted into the DNA stack (143).

4. AAG (Alkyladenine DNA glycosylase)

3-mA DNA glycosylases from different origins such as *E. coli* (AlkA), yeast (MAG), rat (APDG), human (AAG/ANPG) are known as human alkylbase DNA glycosylase that can also excise hypoxanthine. Other human enzymes such as AGT, hABH2, and hABH3 (human AlkB homologs2 and 3) recognize different types of alkylation damage. AlkA has no preference for hypoxanthine pair with different base, whereas mammalian enzymes exhibit a strong preference for I/T and I/G base pairs (142). In *E. coli* and *S. cerevisiae* mutants lacking AAG are very sensitive to killing but not mutagenesis by methylating agents, whereas in *aag*–/– knockout cells, it had been shown that different cell type has various response to alkylating agent. For example, *aag*–/– embryonic stem cells are hypersensitive to simple alkylating agents, *aag*–/– neurons have a low sensitivity, *aag*–/– mouse primary embryonic fibroblasts have either low or no sensitivity, moreover, *aag*–/– myeloid progenitor cells are resistant to alkylation (144).

B. Endonuclease V-mediated repair pathway

Enonuclease V, called initially as deoxyinosine 3' endonuclease, is different from N-glycosylases. This enzyme has no N-glycosylase activity and hydrolyzes the second phosphodiester bond 3' to deoxyinosine or an AP site, generating a nick with a 3' hydroxyl which means the lesion still remains on the DNA in the nicked duplex. In comparision to AlkA, deoxyinosine 3' endonuclease has no preference and exhibited 25%

activity on a single-stranded oligonucleotide substrate containing deoxyinosine whereas AlkA has preference on T/I>G/I>>C/I, A/I. In comparision to AP endonuclease, typical AP ndonucleases cleave the phosphodiester bond immediately 3' or 5' to the AP site. Here, we will discuss endoneclease V from *E. coli*, *Thermotoga maritima* and mammals.

a. Endonuclease V in *E. coli*

E. coli Endo V was initially purified as a novel hypoxanthine-specific endonuclease**,** deoxyinosine 3'-endonuclease (145). The 25-kDa protein recognizes hypoxanthine in both double- and single-stranded DNA. In contrast to the hypoxanthine DNA N-glycosylase, deoxyinosine 3'-endonuclease excises the DNA at the second phosphodiester bond 3' to deoxyinosine, leaving behind a nick with 3'-OH group and 5'-phosphate. The gene coding for endonuclease V (*nfi* for endonuclease V) was found to be identical to orf 225, located at 90 minutes of the *Escherichia coli* genome (GenBankTM ascension no. U00006) (146*).* Endo V has broad range of substrates by recognizing several well defined DNA lesions, including deoxyinosine, urea residues, AP sites, base mismatches, loops and hairpins, flaps, pseudo-Y DNA structures and uracil.

Endo V requires divalent metals to facilitate its activity, such as $Mg2^{+}$, $Co2^{+}$ and $Mn2^{+}$ (147). *E. coli* nfi mutant has stimulated A:T→G:C transitions and G:C→A:T transitions, however, nfi mutants did not cause high frequency of spontaneous mutation at aerobic growth in a rich medium (73). Unlike the length of the repair patch as either 'short-patch' BER (one nucleotide) or 'long-patch' BER (LP-BER; more than one nucleotide), *E. coli* Endo V has its own repair patch. It was identified that the size of

patch DNA at the endo V cleavage site was that 3 nucleotides were preferentially erased from 3' end and 2 nucleotides were erased from 5'end (149).

b. Endonuclease V in bacteria

Endonucleae V has been conserved in several bacterial, archaeal, and lower eukaryotic genomes including *Schizosaccharomyces pombe*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, rats, mice, and humans (150, 151).

Hyperthermophilic bacterium *Thermotoga maritima* (Tma) can also cleave hypoxanthine, AP site, uracil and mismatches. It cleaves only on the inosine-containing strand when an inosine-containing DNA substrate is in excess; however, interestingly, when the enzyme is in excess, a free *Tma* endo V may bind to the complementary strand and make a second nick, resulting in a double-strand break (150). Site-directed mutagenesis revealed that Y80A, H116A and R88A had reduced affinities for double-stranded or single-stranded inosine substrates or nicked products (152). *Tma* endo V is known to act as an endonuclease, however, it was also identified as an exonuclease with $MnCl_2$ as a cofactor (153).

The endonuclease V homologue from *Archaeoglobus fulgidus* (Afu) can only cleave deoxyinosine-containing DNA (154) whereas *Salmonella typhimurium* (Sty) endo V shows DNA cleavage activities against deoxyinosine, deoxyxanthosine, deoxyuridine and mismatches (151).

c. Endonuclease V in mammals

Mouse endo V has 32% sequence homology with *E. coli* endo V, while mouse endo V possesses DNA repair activities that are similar on hypoxanthine containing

DNA, but more limited (155). Based on this paper (155), the mouse enod V is dependent on Mg2+ and higher activity was detected on single-stranded than on double-stranded substrate with the following substrate preference: ss DNA containing hypoxanthine > ds DNA containing hypoxanthine >>> ss DNA containing uracil, while no significant activity was observed towards uracil residues in double-stranded DNA, nor against 8-oxoguanine, C:C mismatches, AP-sites or 5′ flap structures.

C. Nucleotide excision repair (NER)

NER is a multi-step DNA repair pathway that can remove a broad range of bulky, helix-distorting lesions. NER is responsible for the repair of UV light-induced photoproducts, lipid peroxidation induced DNA adducts, cigarette smoke-induced benzo[a]pyrene DNA adducts, chemical carcinogen-induced 4-nitroquinoline (156), removal of oxidative DNA lesions (157). In *Escherichia coli*, the UvrABC enzyme can repair single strand breaks (SSBs) (158). In mammalian systems, NER is carried out by multi-protein complex in a stepwise manner. First, DNA damage recognition, second, excision of damaged site by complexed proteins, last, synthesis and ligation. Here, we will discuss about step one since this step is consist of two distinct pathways, namely global genome repair (GGR) and transcription-coupled repair (TCR), which are identical except for the mode of the DNA damage recognition.

In GGR, the XPC−hHR23 complex is a major DNA damage recognition factor. XPC−hHR23 can sequentially recruit the transcription factor IIH (TFIIH), XPA and replication protein A (RPA) to form a pre-incision complex in the damage sites. TFIIH consists of nine components, among two helicases XPB and XPD which can unwind the

DNA double helix at the damaged site to form opened DNA complex. The dual incision process is performed by endonuclease XPG that makes 3'-incision approximately hydrolyse phosphodiester bonds 2–8 nucleotides of the damage site and the XPF–ERCC1 complex that makes 5'-incision approximately hydrolyse phosphodiester bonds 15–24 nucleotides of damage site. The resulting gap is filled in by DNA polymerases delta/epsilon (Pol δ/ε) which requires proliferating cell nuclear antigen (PCNA), RPA and replication factor C (RFC). Finally, the DNA fragments are ligated by DNA ligase I (Lig1) (159, 160, 161).

In TCR, the removal of DNA lesions from the transcribed strands is much more efficient that GGR. TCR is initiated by stalled RNA polymerase (RNAP) at the damaged site in the transcribed strand. CSB protein interacts with stalled RNAP and recruits other TCR protein factors to damage site such as XPG, TFIIH, RPA and XPA. After XPF is bound to damage site, dual incision process is occurred by CSB-dependent XPG, XPF endonuclease. Synthesis and ligation step are the same as GGR (159).

D. Mismatch DNA repair (MMR)

Mismatch repair (MMR) is the major post-replicative DNA repair system that increases replication fidelity up to 1000 fold (162). Microsatellite instability (MSI) is an established biomarker for MMR dysfunction in tumor cells (163). MMR can repair modified bases such as 8‑oxoG, 2-oxoadenine, carcinogen adducts, and UV-photoproducts (162).

In *E. coli* MMR system, MutS, MutL, MutH and UvrD were identified in studies of mutator strains as key components (164, 165). Initially, MutS protein dimer (or

tetramer) binds to small insertion/deletion loops (IDLs) and unpaired bases. ATP-dependent MutL protein is recruited in MutS to form the ternary complex of MutS-MutL- mismatch sites. It activates monomeric MutH endonuclease which incises an unmethylated GATC sequence at a site 5' or 3' to the mismatch, located even 1000 bp from the mismatch (162, 166, 167). The nick is regard as an entry point of UvrD helicase and SSB (single stranded DNA binding) protein. If the incision occurred 5' to the mismatch, the ssDNA flap is degraded in the 5'→ 3' direction by Exo VIII / RecJ exonuclease or if the incision occurred 3' to the mismatch, it is degraded in the the 3'→ 5' direction by ExoI, ExoVII or Exo X exonuclease. Finally, the SSB-stabilized single-stranded gap is filled in by DNA polymerase III holoenzyme and DNA ends are sealed by LigI (162, 168, 169).

In mammalian MMR system, several homologues of MutS, Mut L are identified in mammalian such as five MutS (MSH2–MSH6) and four MutL (MLH1, MLH3, PMS1, and PMS2) homologues (162). MSH2-MSH6 form a heterodimer, termed MutSα, which recognizes mismatch sites up to about 10 unpaired nucleotides, whereas the MSH2-MSH3 heterodimer (MutSβ) recognizes from 2 up to 16 nucleotides. Once MutSα recruites MutLα, MutSα–MutLα complexes may travel along the DNA helix to recognize mismatch site with interacting to ExoI, PCNA and high mobility group box 1 (HMGB1) protein, which was shown to interact with MutSα (162, 170). When a single-strand break is localized at the 5' side of the mismatch, Exo1, stimulated by MutSα hydrolyzes DNA in the 5'→ 3' direction in replication protein A (RPA)-dependent manner whereas single-strand break located at the 3' side of the mismatch, 3'→5' excision is mediated by Pol δ

or Pol ε proofreading activity (162, 171). Recently, one group has discovered that MutLα has an endonuclease activity, stimulated by MMR cofactors (MutSα, MutLα, PCNA, RFC, ATP and divalent cations) which is revelent to explain about 3' side of the mismatch (172). Lastly, DNA polymerase δ fills the gap, in the presence of PCNA and RPA and followed by seals the ends DNA ligase (probably Lig1) (162).

E. Double strand break (DSB) repair

The simplest mechanism to repair of DSB is nonhomologous DNA end joining (NHEJ). This repair pathway rejoins blunt ends in a manner that does not require to be error free, however, it can lead to sequence modification at the break point when the ends are not compatible. Ku70, Ku80, DNA-dependent protein kinase catalytic subunit (DNA-PKcs), XRCC4, XLF, and DNA ligase IV are required for NHEJ as a core component (173).

Homologous recombition (HR) repair is directed by longer stretches of homology, generally more than 100 bp and is accurate pathway because the undamaged sister chromatid is used as a repair template. Whereas NHEJ can function throughout the cell cycle, HR is largely restricted to late S/G2 phases (174). MRE11, RAD50, NSB1, phosphorylated-histone-H2AX, RAD51, BRCA2, RAD52, RAD54, RAD54B, XRCC2, and XRCC3 are involved in HR (162, 175, 176).

V. DNA repair deficiency-related diseases

If the DNA-repair pathways are not properly processed or acquired defect, it will cause a number of disorders or syndromes such as neurologic symptoms (xeroderma

pigmentosum, Cockayne syndrome, ataxia–telangiectasia, Nijmegen breakage syndrome and Alzheimer disease), cancer (HNPCC, breast cancer and prostate cancer), and accelerated aging (*i.e.*, Bloom syndrome, Rothmund–Thomson syndrome, Werner syndrome) (162, 177, 148).

Here, I will briefly discuss about what defect repair pathway might be involved in disorders or symptons. SCID (severe combined immunodeficiency) is resulted from defecting in V(D)J recombination by mutation in LIG4 and Artemis (63,56) and Alzheimer disease can be caused by defect of two DNA-repair pathways (76). Fanconi anemia (FA) is a rare recessive disease and there are at least 13 genes whose mutations are known to cause FA such as FANCA, FANCB, FANCC, FANCD1, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCJ, FANCL, FANCM and FANCN. Numbers of studied had been on FA, it revealed that mostly NER and NHEJ related protein are involved in. Werner syndrome (WS) is an autosomal recessive progeroid disease involving the mutation of WRN gene belongs to the RecQ helicase family. Moreover, the human genome contains four other RecQ helicase family members, RecQ1, BLM, RecQ4L, and RecQ5. Mutations in BLM and RecQ4L cause Bloom syndrome (BS) known as Bloom–Torre–Machacek syndrome and Rothmund–Thompson syndrome (RTS), respectively. Those three symptons are highly related to cancer. Ataxia–telangiectasia (AT), Nijmegen breakage syndrome (NBS), and ataxia–telangiectasia-like disorder (ATLD) are thought to be chromosome-instability disorders, and the associated defective genes are ATM, NBS1 and MRE11, respectively. Xeroderma pigmentosum (XP), trichothiodystrophy (TTD) and Cockayne syndrome (CS) have mutated in the NER

pathway, which involves more than 20 genes. Seven complementation groups (XPA, ERCC3/XPB, XPC, ERCC2/XPD, DDB1/XPE, ERCC4/XPF, and ERCC5/XPG) are related to mutations in seven genes that play a role in the NER pathway. In CS, CSA (ERCC8) and CSB (ERCC6) genes have been mutated and should be required for transcription-coupled repair. In TTD, Mutations in three genes, XPD (ERCC2), XPB (ERCC3), and TTD-A. Lynch syndrome (HNPCC or Hereditary nonpolyposis colorectal cancer) is an autosomal dominant genetic condition which has a failure of MMR pathway such as MSH2, MLH1, MSH6, PMS1, PMS2, MLH3, and EXO1. Breast cancer is associated with germline mutations in ten different genes such as BRCA1, BRCA2, p53, PTEN, CHEK2, ATM, NBS1, RAD50, BRIP1, and PALB2 and prostate cancer is associated with BRCA1/2, OGG1, XRCC1, CHEK2, and ADPRT (177).

VI. References

1. Pauly N, Pucciariello C, Mandon K, Innocenti G, Jamet A, Baudouin E, Hérouart D, Frendo P, Puppo A. Reactive oxygen and nitrogen species and glutathione: key players in the legume-Rhizobium symbiosis, J Exp Bot. 57 (2006) 1769-76.

2. Dro¨ge, W. Oxidative stress and ageing: is ageing a cysteine deficiency syndrome?, Phil. Trans. R. Soc. B 360 (2005) 2355–72.

3. Hernández-García D, Wood CD, Castro-Obregón S, Covarrubias L. Reactive oxygen species: A radical role in development?, Free Radic Biol Med. (2010) [Epub ahead of print].

4. Fridovich, I. Mitochondria: are they the seat of senescence?, Aging Cell 3 (2004) 13–16.

5. Forman HJ, Fukuto JM, Torres M. Redox signaling: thiol chemistry defines which reactive oxygen and nitrogen species can act as second messengers. Am J Physiol Cell Physiol. 287 (2004) 246-56.

6. Nathan C. Specificity of a third kind: reactive oxygen and nitrogen intermediates in cell signaling. J Clin Invest. 111 (2003) 769–78.

7. Martínez MC, Andriantsitohaina R. Reactive nitrogen species: molecular mechanisms and potential significance in health and disease. Antioxid Redox Signal. 11 (2009) 669-702.

8. Nathan C and Xie QW. Regulation of biosynthesis of nitric oxide. J Biol Chem. 269 (1994) 13725–28.

9. Liu VWT and Huang PL. Cardiovascular roles of nitric oxide: a review of insights from nitric oxide synthase gene disrupted mice. Cardiovasc Res. 77 (2008) 19–29.

10. Nordberg J, Arnér ES. Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. Free Radic Biol Med. 31 (2001) 1287-312.

11. Holmgren A. Thioredoxin. Annu Rev Biochem. 54 (1985) 237–71.

12. Tagaya Y, Maeda Y, Mitsui A, Kondo N, Matsui H, Hamuro J, Brown N, Arai K, Yokota T, Wakasugi H, and et al. ATLderived factor (ADF), an IL-2 receptor/Tac inducer homologous to thioredoxin; possible involvement of dithiol-reduction in the IL-2 receptor induction. EMBO J. 8 (1989) 757–64.

13. Yodoi J and Uchiyama T. Diseases associated with HTLVI: virus, IL-2 receptor dysregulation and redox regulation. Immunol Today. 13(1992) 405–11.

14. Spyrou G, Enmark E, Miranda–Vizuete A, and Gustafsson J. Cloning and expression of a novel mammalian thioredoxin. J Biol Chem. 272 (1997) 2936–41.

15. Hirota K, Nakamura H, Masutani H, and Yodoi J. Thioredoxin superfamily and thioredoxin-inducing agents. Ann NY Acad Sci. 957(2002) 189–89.

16. Nakamura H. Thioredoxin and its related molecules: update 2005. Antioxid Redox Signal 7(2005) 823–28.

17. Kondo N, Nakamura H, Masutani H, Yodoi J. Redox regulation of human thioredoxin network. Antioxid Redox Signal. 8 (2006) 1881-90.

18. Holmgren A. Thioredoxin and glutaredoxin systems. J Biol Chem. 264 (1989) 13963–66.

19. Holmgren A, Bjo¨rnstedt M. Thioredoxin and thioredoxin reductase. Methods Enzymol. 252 (1995) 199–208.

20. Gleason FK, Holmgren A. Thioredoxin and related proteins in procaryotes. FEMS Microbiol. Rev. 4 (1988) 271–97.

21. Arne´r ES, Zhong L, Holmgren A. Preparation and assay of mammalian thioredoxin and thioredoxin reductase. Methods Enzymol. 300 (1999) 226–39.

22. Watson WH, Yang X, Choi YE, Jones DP, Kehrer JP. Thioredoxin and its role in toxicology. Toxicol Sci. 78 (2004) 3-14.

23. Spector A, Yan GZ, Huang RR, McDermott MJ, Gascoyne PR, Pigiet V. The effect of H2O2 upon thioredoxin-enriched lens epithelial cells. J Biol Chem. 263 (1988) 4984-90.

24. Ritz D, Patel H, Doan B, Zheng M, Aslund F, Storz G, Beckwith J. Thioredoxin 2 is involved in the oxidative stress response in Escherichia coli. J Biol Chem. 275 (2000) 2505-12.

25. Powis, G., Gasdaska, J. R., and Baker, A. Redox signaling and the control of cell growth and death. Adv Pharmacol. 38 (1997) 329–59.

26. Andoh, T., Chock, P. B., and Chiueh, C. C. The roles of thioredoxin in protection against oxidative stress-induced apoptosis in SH-SY5Y cells. J Biol Chem. 277 (2002) 9655–60.

27. Shioji, K., Kishimoto, C., Nakamura, H., Masutani, H., Yuan, Z., Oka, S-i, and Yodoi, J. Overexpression of thioredoxin-1 in transgenic mice attenuates adriamycin-induced cardiotoxicity. Circulation 106 (2002) 1403–09.

28. Tanaka, T., Nakamura, H., Nishiyama, A., Hosoi, F., Masutani, H., Wada, H. and Yodoi, J. Redox regulation by thioredoxin superfamily; protection against oxidative stress and aging. Free Radic Res. 33 (2000) 851–55.

29. Nishiyama, A., Masutani, H., Nakamura, H., Nishinaka, Y., and Yodoi, J. Redox regulation by thioredoxin and thioredoxin-binding proteins. IUBMB Life 52 (2001) 29–33.

30. Sun J, Steenbergen C, Murphy E. S-nitrosylation: NO-related redox signaling to protect against oxidative stress. Antioxid Redox Signal. 8 (2006) 1693-1705.

31. Keshive, M., Singh, S., Wishnok, J. S., Tannenbaum, S. R., and Deen, W. M. Kinetics of S-nitrosation of thiols in nitric oxide solutions. Chem Res Toxicol. 9 (1996) 988–993.

32. Wink, D. A., Nims, R. W., Darbyshire, J. F., Christodoulou,D., Hanbauer, I., Cox, G. W., Laval, F., Laval, J.,Cook, J. A., and Krishna, M. C. Reaction kinetics for nitrosation of cysteine and glutathione in aerobic nitric oxide solutions at neutral pH. Insights into the fate and physiological effects of intermediates generated in the NO/O2 reaction. Chem Res Toxicol. 7 (1994) 519–525.

33. Mitchell, D. A., Morton, S. U., and Marletta, M. A. Thioredoxin catalyzes the S-nitrosation of the caspase-3 active site cysteine. Nat Chem Biol. 1 (2005) 154–158.

34. Haendeler J, Hoffmann J, Tischler V, Berk BC, Zeiher AM, and Dimmeler S. Redox regulatory and anti-apoptotic functions of thioredoxin depend on S-nitrosylation at cysteine 69. Nat Cell Biol. 4 (2002) 743–749.

35. Foster MW and Stamler JS. New insights into protein S-nitrosylation: mitochondria as a model system. J Biol Chem. 279 (2004) 25891–97.

36. Asahi M, Fujii J, Suzuki K, Seo HG, Kuzuya T, Hori M, Tada M, Fujii S, and Taniguchi N. Inactivation of glutathione peroxidase by nitric oxide donor. J Biol Chem. 270 (1995) 21035–39.

37. Butzer U, Weidenbach H, Gansauge S, Gansauge F, Beger HG, and Nussler AK. Increased oxidative stress in the RAW 264.7 macrophage cell line is partially mediated via the S-nitrosothiol-induced inhibition of glutathione reductase. FEBS Lett. 445 (1999) 274–78.

38. Lo Bello M, Nuccetelli M, Caccuri AM, Stella L, Parker MW, Rossjohn J, McKinstry WJ, Mozzi AF, Federici G, Polizio F, Pedersen JZ, and Ricci G. Human glutathione transferase P1–1 and nitric oxide carriers: a new role for an old enzyme. J Biol Chem. 276 (2001) 42138–45.

39. Yasinska IM, Kozhukhar AV, and Sumbayev VV. S-nitrosation of thioredoxin in the nitrogen monoxide/superoxide system activates apoptosis signal-regulating kinase 1. Arch Biochem Biophys. 428 (2004) 198–203.

40. Jackson AL, Loeb LA. The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. Mutat Res. 477 (2001) 7-21.

41. Krwawicz J, Arczewska KD, Speina E, Maciejewska A, Grzesiuk E. Bacterial DNA repair genes and their eukaryotic homologues: 1. Mutations in genes involved in base excision repair (BER) and DNA-end processors and their implication in mutagenesis and human disease. Acta Biochim Pol. 54 (2007) 413-34.

42. Valko M, Rhodes CJ, Moncol J, Izakovic M, Mazur M. Free radicals, metals and antioxidants in oxidative stressinducedm cancer. Chem Biol Interact 160 (2006) 1-40.

43. Halliwell B, Aruoma OI. DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems. FEBS Lett. 281 (1991) 9-19.

44. Altieri F, Grillo C, Maceroni M, Chichiarelli S. DNA damage and repair: from molecular mechanisms to health implications. Antioxid Redox Signal. 10 (2008) 891-937.

45. Lindahl T. Instability and decay of the primary structure of DNA. Nature 362 (1993) 709-15.

46. Halliwell B. Oxygen and nitrogen are pro-carcinogens. Damage to DNA by reactive oxygen, chlorine and nitrogen species: measurement, mechanism and the effects of nutrition. Mutat Res. 443 (1999) 37-52.

47. Kreutzer DA, Essigmann JM. Oxidized, deaminated cytosines are a source of C-->T transitions *in vivo*. Proc Natl Acad Sci USA 95 (1998) 3578-82.

48. Purmal AA, Lampman GW, Bond JP, Hatahet Z, Wallace SS. Enzymatic processing of uracil glycol, a major oxidative product of DNA cytosine. J Biol Chem. 273 (1998) 10026-35.

49. Kavli B, Otterlei M, Slupphaug G, Krokan HE. Uracil in DNA-general mutagen, but normal intermediate in acquired immunity. DNA Repair (Amst) 6 (2007) 505-16.

50. Kaudewitz P. Production of bacterial mutants with nitrous acid. Nature 183 (1959) 1829–30.

51. Vielmetter W, Schusler H. The base specificity of mutations induced by nitrous acid. Biochem Bioph Res Co 1 (1960) 324–8.

52. Parry TE, Blackmore JA. Serum uracil and uridine levels in normal subjects and their possible significance. J Clin Pathol 24 (1974) 789–93.

53. Parry TE. The non-random distribution of point mutations in leukaemia and myelodysplasia: a possible pointer to their aetiology. Leuk Res 21(1997) 539–74.

54. Sousa MM, Krokan HE, Slupphaug G. DNA-uracil and human pathology. Mol Aspects Med. 28 (2007) 276-306.

55. Shapiro, R. Damage to DNA caused by hydrolysis. In: Kleppe, E.S.a.K. (Ed.), NATO Advanced Study Institute Series, Life Sciences. Plenum Press, New York and London 40 (1981) 3–18.

56. Moshous D, Callebaut I, de Chasseval R, Corneo B, Cavazzana-Calvo M, Le Deist F, Tezcan I, Sanal O, Bertrand Y, Philippe N, Fischer A, de Villartay JP. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human evere combined immune deficiency. Cell 105 (2001) 177–186.

57. Ehrlich, M., and R. Y.-H. Wang. 5-Methylcytosine in eukaryotic DNA. Science 212 (1981) 1350-57.

58. Gama-Sosa, M. A., R. Midgett, V. A. Slagel, S. Githens, K. C. Kuo, C. W. Gehrke, and M. Ehrlich. Tissue-specific differences in DNA methylation in various mammals. Biochim. Biophys. Acta 740 (1983) 212-19.

59. Cooper, D. N. Eukaryotic DNA methylation. Hum. Genet. 64 (1983) 315-33.

60. Kuo, K., R. A. McCune, C. Gehrke, R. Midgetts, and M. Ehrlich. Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleoside in DNA. Nucleic Acids Res. 8 (1980) 4763-76.

61. Parry TE. On the mutagenic action of adenine. Leuk Res. 31 (2007) 1621-24.

62. Kaudewitz P. Production of bacterial mutants with nitrous acid. Nature 183(1959) 1829–30.

63. O'Driscoll M, Gennery AR, Seidel J, Concannon P, and Jeggo PA. An overview of three new disorders associated with genetic instability: LIG4 syndrome, RS-CID and ATRSeckel syndrome. DNA Repair (Amst) 3 (2004) 1227–35.

64. Lindahl, T**.** DNA Glycosylases, Endonucleases for Apurinic/Apyrimidinic Sites and Base Excision Repair. Prog. Nucleic Acid Res. Mol. Biol. 22 (1979) 135-192.

65. Bradshaw, J. S., and Kuzminov, A. RdgB acts to avoid chromosome fragmentation in Escherichia coli. Mol Microbiol 48 (2003) 1711-25.

66. Suzuki T., Matsumura,Y., Ide,H., Kanaori,K., Tajima,K. and Makino,K. Deglycosylation susceptibility and base-pairing stability of 2′-deoxyoxanosine in oligodeoxynucleotide. Biochemistry 36 (1997) 8013–19.

67. Lindahl T. and Nyberg,B. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. Biochemistry 13 (1974) 3405–10.

68. Shapiro R. and Klein,R.S. The deamination of cytidine and cytosine by acidic buffer solutions. Mutagenic implications. Biochemistry 5 (1966) 2358–62.

69. Caulfield J.L., Wishnok,J.S. and Tannenbaum,S.R. Nitric oxide-induced deamination of cytosine and guanine in deoxynucleosides and oligonucleotides. J. Biol. Chem. 273 (1998) 12689–95.

70. Nguyen T., Brunson,D., Crespi,C.L., Penman,B.W., Wishnok,J.S. and Tannenbaum,S.R. DNA damage and mutation in human cells exposed to nitric oxide *in vitro*. Proc Natl Acad Sci. USA, 89 (1992) 3030–34.

71. Eritja R., Horowitz,D.M., Walker,P.A., Ziehler-Martin,J.P., Boosalis,M.S., Goodman,M.F., Itakura,K. and Kaplan,B.E. Synthesis and properties of oligonucleotides containing 2′-deoxynebularine and 2′-deoxyxanthosine. Nucleic Acids Res. 14 (1986) 8135–53.

72. Kamiya H., Miura,H., Suzuki,M., Murata,N., Ishikawa,H., Shimizu,M., Komatsu,Y., Murata,T., Sasaki,T., Inoue,H. Mutations induced by DNA lesions in hot spots of the c-Ha-ras gene. Nucleic Acids Symp. Ser. 1 (1992) 179–180.

73. Schouten K.A. and Weiss,B. Endonuclease V protects *Escherichia coli* against specific mutations caused by nitrous acid. Mutat Res 435 (1999) 245–54.

74. He B., Qing,H. and Kow,Y.W. Deoxyxanthosine in DNA is repaired by *Escherichia coli* endonuclease V. Mutat Res 459(2000) 109–14.

75. Friedkin, M. Enzymatic Synthesis of Desoxyxanthosine by the Action of Xanthosine Phosphorylase in Mammalian Tissue. J. Am. Chem. Soc.74 (1952) 112−15.

76. Fishel ML, Vasko MR, and Kelley MR. DNA repair in neurons: so if they don't divide what's to repair? Mutat Res 614 (2007) 24–36.

77. Wuenschell GE, O'Connor TR, Termini J. Stability, miscoding potential, and repair of 2'-deoxyxanthosine in DNA: implications for nitric oxide-induced mutagenesis. Biochemistry. 42 (2003) 3608-16.

78. Suzuki, T., Yoshida, M., Yamada, M., Ide, H., Kobayashi, M., Kanaori, K., Tajima, K., and Makino, K. Misincorporation of 2'-deoxyoxanosine 5'-triphosphate by DNA polymerases and its implication for mutagenesis. Biochemistry 37 (1998) 11592-98.

79. Suzuki, T., Yamada, M., Ide, H., Kanaori, K., Tajima, K., Morii, T., and Makino, K. Identification and characterization of a reaction product of 2'-deoxyoxanosine with glycine. Chem. Res. Toxicol. 13 (2000) 227-30.

80. Hitchcock TM, Dong L, Connor EE, Meira LB, Samson LD, Wyatt MD, Cao W. Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase. J Biol Chem. 279 (2004) 38177-83.

81. Shimada N., Yagisawa,N., Naganawa,H., Takita,T., Hamada,M., Takeuchi,T. and Umezawa,H. Oxanosine, a novel nucleoside from Actinomycetes. J. Antibiot. 34 (1981) 1216–18.

82. Yagisawa N., Shimada,N., Takita,T., Ishizuka,M., Takeuchi,T. and Umezawa,H. Mode of action of oxanosine, a novel nucleoside antibiotic. J. Antibiot. 35 (1985) 755–59.

83. Hernández B, Soliva R, Luque FJ, Orozco M. Misincorporation of 2'-deoxyoxanosine into DNA: a molecular basis for NO-induced mutagenesis derived from theoretical calculations. Nucleic Acids Res.28 (2000) 4873-83.

84. Hazra TK, Das A, Das S, Choudhury S, Kow YW, and Roy R. Oxidative DNA damage repair in mammalian cells: a new perspective. DNA Repair (Amst) 6 (2007) 470–480.

85. Hitomi K, Iwai S, and Tainer JA. The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. DNA Repair (Amst) 6 (2007) 410–428.

86. Sung JS and Demple B. Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. FEBS J 273 (2006) 1620–29.

87. Wilson III DM and Bohr VA. The mechanics of base excision repair, and its relationship to aging and disease. DNA Repair (Amst) 6 (2007) 544–559.

88. Huffman JL, Sundheim O, and Tainer JA. DNA base damage recognition and removal: new twists and grooves. Mutat Res 577(2005) 55–76.

89. Radman,M. An endonuclease from Escherichia coli that introduces single polynucleotide chain scissions in ultravioletirradiated DNA. J Biol Chem 251 (1976) 1438-45.

90. Jorgensen,T.J., Kow,Y.-W., Wallace,S.S. and Henner,W.D. Mechanism of action of Micrococcus luteus y-endonuclease. Biochemistry 26 (1987) 6436-43.

91. Gosset,J., Lee,K., Cunningham,R.P. and Doetsch,P.W. Yeast redoxy endonuclease, a DNA repair enzyme similar to Escherichia coli endonuclease III. Biochemistry 27 (1988) 2629-34.

92. Doetsch,P.W., Henner,W.D., Cunningham,R.P., Toney,J.H. and Helland,D.E. A highly conserved endonuclease activity present in Escherichia coli, bovine and human cells recognizes oxidative DNA damage at sites of pyimidines. Mol Cell Biol 7 (1987) 26-32.

93. Higgins,S.A., Frenkel,K., Cummings,A. and Teebor,G.W. Definitive characteri zation of human thymine glycol N-glycosylase activity. Biochemistry 26 (1987) 1683-88.

94. Ganguly,T., Weems,K.M. and Duker,N.J. Ultraviolet-induced thymine hydrates in DNA are excised by bacterial and human DNA glycosylase activities. Biochemistry 29 (1990) 7222-28.

95. Michaels,M.L., Pham,L., Nghiem,Y., Cruz,C. and Miller,J.H. MutY, an adenine glycosylase active on G-A mispairs, has homology to endonuclease III. Nucleic Acids Res 18 (1990) 3841-45.

96. Wallace, S. DNA Damages Processed by Base Excision Repair: Biological Consequences. Int. J Radiat Biol 66 (1994) 579-89.

97. Katafuchi A, Nakano T, Masaoka A, Terato H, Iwai S, Hanaoka, Ide H. Differential specificity of human and Escherichia coli endonuclease III and VIII homologues for oxidative base lesions. J Biol Chem 279 (2004) 14464–71.

98. Miller H, Fernandes AS, Zaika E, McTigue MM, Torres MC, Wente M, Iden CR, Grollman AP. Stereoselective excision of thymine glycol from oxidatively damaged DNA. Nucleic Acids Res 32 (2004) 338–45.

99. Nash HM, Bruner SD, Schaerer OD, Kawate T, Addona TA, Spooner E, Lane WS & Verdine GL. Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a baseexcision DNA-repair protein superfamily. Curr Biol 6 (1996) 968–980.

100. Van der Kemp PA, Thomas D, Barbey R, de Oliveira R & Boiteux S. Cloning and expression in Escherichia coli of the OGG1 gene of Saccharomyces cerevisiae, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. P Natl Acad Sci USA 93(1996) 5197–202.

101. Klungland A, Laerdahl JK & Rognes T. OGG1: from structural analysis to the knockout mouse. Oxidative Damage to Nucleic Acids (Evans MD & Cooke MS, eds) 2007 67–80.

102. Jensen A, Calvayrac G, Karahalil B, Bohr VA & Stevnsner T. Mammalian 8-oxoguanine DNA glycosylase 1 incises 8-oxoadenine opposite cytosine in nuclei and mitochondria,while a different glycosylase incises 8-oxoadenine opposite guanine in nuclei. J Biol Chem 278 (2003) 19541–48.

103. Sartori AA, Lingaraju GM, Hunziker P, Winkler FK & Jiricny J.Pa-AGOG, the founding member of a new family of archaeal 8-oxoguanine DNA-glycosylases. Nucleic Acids Res 32 (2004) 6531–39.

104. T. Lindahl, An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues. Proc Natl Acad Sci U. S. A. 71 (1974) 3649–53.

105. H. Nilsen, M. Otterlei, T. Haug, K. Solum, T.A. Nagelhus, F. Skorpen, H.E. Krokan. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. Nucl Acids Res 25 (1997) 750–55.

106. Nagelhus, T. A., Haug, T., Singh, K. K., Keshav, K. F., Skorpen, F., Otterlei, M., Bharati, S., Lindmo, T., Benichou, S., Benarous, R., and Krokan, H. E. A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A. J Biol Chem 272 (1997) 6561-66.

107. Nilsen, H., Rosewell, I., Robins, P., Skjelbred, C.E., Andersen, S., Slupphaug, G.,Daly, G., Krokan, H.E., Lindahl, T. and Barnes, D.E. Mol Cell (2000) in press.

108. G. Slupphaug, L.C. Olsen, D. Helland, R. Aasland and H.E. Krokan, Cell cycle regulation and in vitro hybrid arrest analysis of the major human uracil-DNA glycosylase. Nucleic Acids Res. 19 (1991) 5131–37

109. R. Savva, K. McAuley-Hecht, T. Brown and L. Pearl, The structural basis of specific base-excision repair by uracil-DNA glycosylase. Nature 373 (1995) 487–93.

110. C.D. Mol, A.S. Arvai, G. Slupphaug, B. Kavli, I. Alseth, H.E. Krokan and J.A. Tainer, Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. Cell 80 (1995) 869–78

111. C.D. Putnam, M.J. Shroyer, A.J. Lundquist, C.D. Mol, A.S. Arvai, D.W. Mosbaugh and J.A. Tainer, Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with Escherichia coli uracil-DNA glycosylase. J Mol Biol 287 (1999) 331–46

112. S.S. Parikh, C.D. Mol, G. Slupphaug, S. Bharati, H.E. Krokan and J.A. Tainer, Base-excision repair initiation revealed by crystal structures and DNA-binding kinetics of human uracil-DNA glycosylase bound to DNA. EMBO J. 17 (1998) 5412–26.

113. P. Nedderman, J. Jiricny, The purification of a mismatchspecific thymine-DNA glycosylase from HeLa cells. J Biol Chem 268 (1993) 21218–24.

114. P. Nedderman, J. Jiricny, Efficient removal of uracil fromG–U mispairs by the mismatch-specific thymine DNA glycosylasefrom HeLa-cells. Proc Natl Acad Sci USA 91 (1994) 1642–46.

115. P. Gallinari, J. Jiricny. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. Nature 383 (1996) 735–38.

116. T.E. Barrett, R. Savva, G. Panayotou, T. Barlow, T. Brown,J. Jiricny, L.H. Pearl. Crystal structure of a G:TrU mismatch mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interaction. Cell 92 (1998) 117–29.

117. Métivier R., Gallais R., Tiffoche C., Le Péron C., Jurkowska R. Z., Carmouche R. P., Ibberson D., Barath P., Demay F., Reid G., Benes V., Jeltsch A., Gannon F., Salbert G. Cyclical DNA methylation of a transcriptionally active promoter. Nature 452 (280) 45–50.

118. Kangaspeska S., Stride B., Metivier R., Polycarpou-Schwarz M., Ibberson D., Carmouche R. P., Benes V., Gannon F., Reid G. Transient cyclical methylation of promoter DNA. Nature 452 (2008) 112–15.

119. Maiti A., Morgan M. T., Pozharski E., Drohat A. C. Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 8890–95.

120. Maiti A, Morgan MT, Drohat AC. Role of two strictly conserved residues in nucleotide flipping and N-glycosylic bond cleavage by human thymine DNA glycosylase. J Biol Chem. 284 (2009) 36680-8.

121. Smet-Nocca C, Wieruszeski JM, Chaar V, Leroy A, Benecke A. The thymine-DNA glycosylase regulatory domain: residual structure and DNA binding. Biochemistry. 47 (2009) 6519-30.

122. Mohan RD, Litchfield DW, Torchia J, Tini M. Opposing regulatory roles of phosphorylation and acetylation in DNA mispair processing by thymine DNA glycosylase. Nucleic Acids Res. 38(2010) 1135-48.

123. Zhu B, Benjamin D, Zheng Y, Angliker H, Thiry S, Siegmann M, Jost JP. Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. Proc Natl Acad Sci U S A. 98 (2001) 5031-6.

124. Tini M, Benecke A, Um SJ, Torchia J, Evans RM, Chambon P. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol. Cell. 9(2002) 265–7.

125. Hardeland U, Steinacher R, Jiricny J, Schar P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J.21(2002) 1456–64.

126. B. Kavli, O. Sundheim, M. Akbari, M. Otterlei, H. Nilsen, F. Skorpen, P.A. Aas, L. Hagen, H.E. Krokan, G. Slupphaug, hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, andUin single-stranded DNA, with hSMUG1 as a broad specificity backup. J. Biol. Chem. 277 (2002) 39926–36.

127. M. Matsubara, A. Masaoka, T. Tanaka, T. Miyano, N. Kato, H. Terato, Y. Ohyama, S. Iwai, H. Ide, Mammalian 5-formyluracil-DNA glycosylase. 1. Identification and characterization of a novel activity that releases 5-formyluracil from DNA. Biochemistry 42 (2003) 4993–5002.

128. R.J. Boorstein, A. Cummings Jr., D.R. Marenstein, M.K. Chan,Y. Ma, T.A. Neubert, S.M. Brown, G.W. Teebor, Definitive identification of mammalian 5-hydroxymethyluracilDNA N-glycosylase activity as SMUG1. J. Biol. Chem. 276 (2001) 41991–97.

129. Darwanto A, Theruvathu JA, Sowers JL, Rogstad DK, Pascal T, Goddard W 3rd, Sowers LC. Mechanisms of base selection by human single-stranded selective monofunctional uracil-DNA glycosylase. J Biol Chem. 284 (2009) 15835-46.

130. Mi R, Dong L, Kaulgud T, Hackett KW, Dominy BN, Cao W. Insights from xanthine and uracil DNA glycosylase activities of bacterial and human SMUG1: switching SMUG1 to UDG. J Mol Biol. 385(2009) 761-78.

131. An Q., Robins P., Lindahl T., Barnes D.E. C→T mutagenesis and gamma-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases. EMBO J. 24 (2005) 2205–13.

132. Visnes T, Doseth B, Pettersen HS, Hagen L, Sousa MM, Akbari M, Otterlei M, Kavli B, Slupphaug G, Krokan HE. Uracil in DNA and its processing by different DNA glycosylases. Philos Trans R Soc Lond B Biol Sci. 364 (2009) 563-8.

133. Wibley JE, Waters TR, Haushalter K, Verdine GL, Pearl LH. Structure and specificity of the vertebrate anti-mutator uracil-DNA glycosylase SMUG1. Mol Cell. 11 (2003) 1647-59.

134. Pettersen HS, Sundheim O, Gilljam KM, Slupphaug G, Krokan HE, Kavli B. Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. Nucleic Acids Res. 35 (2007) 3879-92.

135. Kosaka H, Hoseki J, Nakagawa N, Kuramitsu S, Masui R. Crystal structure of family 5 uracil-DNA glycosylase bound to DNA. J Mol Biol. 373 (2007) 839-50.

136. Sartori, A. A., Fitz-Gibbon, S., Yang, H., Miller, J. H., and Jiricny, J. A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. EMBO J. 21 (2002) 3182-91.

137. Sandigursky M, Franklin WA. Thermostable uracil-DNA glycosylase from Thermotoga maritima a member of a novel class of DNA repair enzymes. Curr Biol. 9 (1999) 531-34.

138. Sandigursky M, Franklin WA. Uracil-DNA glycosylase in the extreme thermophile Archaeoglobus fulgidus. J Biol Chem. 275(2000) 19146-49.

139. Sartori AA, Schär P, Fitz-Gibbon S, Miller JH, Jiricny J. Biochemical characterization of uracil processing activities in the hyperthermophilic archaeon Pyrobaculum aerophilum. J Biol Chem. 276 (2001) 29979-86.

140. Hoseki J, Okamoto A, Masui R, Shibata T, Inoue Y, Yokoyama S, Kuramitsu S. Crystal structure of a family 4 uracil-DNA glycosylase from Thermus thermophilus HB8. J Mol Biol. 333(2003) 515-26.

141. Labahn J, Scharer OD, Long A, Ezaz-Nikpay K, Verdine GL & Ellenberger TE. Structural basis for the excision repair of alkylation-damaged DNA. Cell 86 (1996) 321–329.

142. O'Brien PJ & Ellenberger T. The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site. J Biol Chem 279 (2004) 26876–84.

143. Hollis T, Ichikawa Y & Ellenberger T. DNA bending and a flip-out mechanism for base excision by the helix–hairpin–helix DNA glycosylase, Escherichia coli AlkA. EMBO J 19 (2000) 758–766.

144. Paik J, Duncan T, Lindahl T, and Sedgwick B. Sensitization of human carcinoma cells to alkylating agents by small interfering RNA suppression of 3-alkyladenine-DNA glycosylase. Cancer Res 65 (2005) 10472-77.

145. Yao M, Hatahet Z, Melamede RJ, Kow YW. Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from Escherichia coli. J Biol Chem. 269 (1994) 16260-8.

146. Blattner F. R., Burland V. D., Plunkett G., III, Sofia H. J., Daniels D. L. Analysis of the *Escherichia coli* genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes. Nucleic Acids Res. 21(1993) 5408–17.

147. Feng, H., Dong, L., and Cao, W. Catalytic mechanism of endonuclease v: a catalytic and regulatory two-metal model. Biochemistry 45 (2006) 10251-59.

148. Perry JJ, Fan L, and Tainer JA. Developing master keys to brain pathology, cancer and aging from the structural biology of proteins controlling reactive oxygen species and DNA repair. Neuroscience 145 (2007) 1280–99.

149. Weiss, B. Removal of deoxyinosine from the Escherichia coli chromosome as studied by oligonucleotide transformation. DNA Repair (Amst) 7 (2008) 205-12.

150. Huang J, Lu J, Barany F, Cao W. Multiple cleavage activities of endonuclease V from Thermotoga maritima: recognition and strand nicking mechanism. Biochemistry. 40 (2000) 8738-48.

151. Feng H, Klutz AM, Cao W. Active site plasticity of endonuclease V from Salmonella typhimurium. Biochemistry. 44 (2005) 675-83.

152. Huang, J., Lu, J., Barany, F., and Cao, W. Mutational analysis of endonuclease V from Thermotoga maritime. Biochemistry 41 (2002) 8342−50.

153. Feng H, Dong L, Klutz AM, Aghaebrahim N, Cao W. Defining amino acid residues involved in DNA-protein interactions and revelation of 3'-exonuclease activity in endonuclease V. Biochemistry 44 (2005) 11486-95.

154. Liu, J., He, B., Qing, H., and Kow, Y. W. A deoxyinosine specific endonuclease from hyperthermophile, Archaeoglobus fulgidus: a homolog of Escherichia coli endonuclease V. Mutat. Res.461 (2000) 169−177.

155. Moe A, Ringvoll J, Nordstrand LM, Eide L, Bjørås M, Seeberg E, Rognes T, Klungland A. Incision at hypoxanthine residues in DNA by a mammalian homologue of the Escherichia coli antimutator enzyme endonuclease V. Nucleic Acids Res. 31 (2003) 3893-900.

156. Kraemer KH, Patronas NJ, Schiffmann R, Brooks BP, TamuraD, Digiovanna JJ. Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: A complex genotype-phenotype relationship. Neuroscience 145 (2007) 1388–96.

157. Brooks PJ. The case for 8,5'-cyclopurine-2'-deoxynucleosides as endogenous DNA lesions that cause neurodegeneration in xeroderma pigmentosum. Neuroscience 145 (2007) 1407–17.

158. Truglio JJ, Croteau DL, Van Houten B, Kisker C. Prokaryotic nucleotide excision repair: the UvrABC system. Chem Rev 106 (2006) 233–52.

159. Maddukuri L, Dudzińska D, Tudek B. Bacterial DNA repair genes and their eukaryotic homologues: 4. The role of nucleotide excision DNA repair (NER) system in mammalian cells. Acta Biochim Pol. 54 (2007) 469-82.

160. Sancar A, Reardon JT. Nucleotide excision repair in E. coli and man. Adv Protein Chem 69 (2004) 43–71.

161. Friedberg EC. How nucleotide excision repair protects against cancer. Nat Rev Cancer 1 (2001) 22–33.

162. Arczewska KD, Kuśmierek JT. Bacterial DNA repair genes and their eukaryotic homologues: 2. Role of bacterial mutator gene homologues in human disease. overview of nucleotide pool sanitization and mismatch repair systems. Acta Biochim Pol. 54 (2007) 435-57.

163. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN, Srivastava S. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 96 (2004) 261–268.

164. Cox EC, Degnen GE, Scheppe ML. Mutator gene studies in Escherichia coli: the mutS gene. Genetics 72 (1972) 551–567.

165. Wagner R Jr, Meselson M. Repair tracts in mismatched DNA heteroduplexes. Proc Natl Acad Sci USA 73 (1976) 4135–4139.

166. Welsh KM, Lu AL, Clark S, Modrich P. Isolation and characterization of the Escherichia coli mutH gene product. J Biol Chem 262 (1987) 15624–29.

167. Bruni R, Martin D, Jiricny J. d(GATC) sequences influence Escherichia coli mismatch repair in a distancedependent manner from positions both upstream and downstream of the mismatch. Nucleic Acids Res 16 (1988) 4875–90.

168. Cooper DL, Lahue RS, Modrich P. Methyl-directed mismatch repair is bidirectional. J Biol Chem 268 (1993) 11823–29.

169. Grilley M, Griffith J, Modrich P. Bidirectional excision in methyl-directed mismatch repair. J Biol Chem 268 (1993) 11830–37.

170. Yuan F, Gu L, Guo S, Wang C, Li GM. Evidence for involvement of HMGB1 protein in human DNA mismatch repair. J Biol Chem 279 (2004) 20935-40.

171. Wang H, Hays JB. Mismatch repair in human nuclear extracts. Time courses and ATP requirements for kinetically distinguishable steps leading to tightly controlled 5' to 3' and aphidicolin-sensitive 3' to 5' mispair-provoked excision. J Biol Chem 277 (2002) 26143–48.

172. Kadyrov FA, Dzantiev L, Constantin N, Modrich P. Endonucleolytic function of MutLalpha in human mismatch repair. Cell 126 (2006) 297–308.

173. Spagnolo L, Rivera-Calzada A, Pearl LH, Llorca O. Three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair. Mol Cell 22 (2006) 511–519.

174. Helleday T, Lo J, van Gent DC, Engelward BP. DNA doublestrand break repair: from mechanistic understanding to cancer treatment. DNA Repair (Amst) 6 (2007) 923–35.

175. Shin DS, Chahwan C, Huffman JL, and Tainer JA. Structure and function of the double strand break repair machinery. DNA Repair (Amst) 3 (2004) 863–73.

176. Wyman C and Kanaar R. DNA double-strand break repair: all's well that ends well. Annu Rev Genet 40 (2006) 363–83.

177. Subba Rao K. Mechanisms of disease: DNA repair defects and neurological disease. Nat Clin Pract Neurol 3 (2007) 162–72.

CHAPTER TWO

INVOLVEMENT OF THIOREDOXIN DOMAIN-CONTAINING 5 IN RESISTANCE TO NITROSATIVE STRESS

I. Abstract

Living organisms are exposed to nitrosative stress mediated by nitric oxide (NO) and its derivatives. Multiple cellular mechanisms may be needed to cope with nitrosative stress. This work takes advantage of a hypersensitive *Escherichia coli* genetic system to identify genes involved in resistance to nitrosative stress in mouse lungs. Mouse thioredoxin domain-containing 5 (mTrx 5) was identified as one of the candidate genes. Its ability to complement the hypersensitive phenotype in an *E. coli* mutant strain was confirmed by genetic analysis. Purified recombinant mouse thioredoxin domain-containing 5 protein reduced DNA damage that is sensitive to cleavage by the deamination repair enzyme endonuclease V, indicating that mTrx 5 may play a role in scavenging the reactive nitrogen species. *E. coli* thioredoxin 1 and thioredoxin 2 proteins also reduced the DNA damage in a similar manner. Deletion of *trxA* (encodes thioredoxin 1) or *trxC* (encodes thioredoxin 2) in *E. coli* resulted in a slightly higher sensitivity to nitrosative stress. On the other hand, deletion of both *trxA* and *trxC* greatly increased its sensitivity to nitrosative stress. Complementation with the mTrx 5 gene rescued the sensitive phenotype of the double deletion mutant. The potential roles that mTrx 5 may play in coping with nitrosative stress are discussed.

II. Introduction

Nitrosative stress is mediated by nitric oxide (NO) and its derivatives such as nitrogen dioxide, dinitrogen trioxide, dinitrogen tetraoxide, nitrite, nitrosothiol and peroxynitrite, which are collectively known as reactive nitrogen species (RNS) or

reactive nitrogen intermediates (RNI) [1, 2]. Humans are exposed to nitrosative stress generated from the living environment, food intake, or endogenous sources. In mammalian systems, NO is derived from L-arginine by NO synthases (NOS). Three NOS enzymes have been found in mammalian genomes, which are nNOS from brain neurons, eNOS from endothelial cells, and iNOS (inducible NOS) from macrophages [3-5]. nNOS and eNOS are expressed constitutively, while iNOS is inducible by lipopolysaccharide (LPS) or inflammatory cytokines. While low levels of NO may serve as an intracellular signaling molecule, high doses of NO generated by inducible NO synthase in phagocytes play an important role in containing invading microbial pathogens [2, 6-8]. As an antimicrobial agent, NO is involved in killing microbial pathogens by inflicting damage to macromolecules such as DNA and protein [2, 4, 7]. In aerobic conditions, NO reacts with oxygen to form nitrous anhydride (N_2O_3), a potent nitrosating reagent capable of deaminating DNA bases [9]. NO can also react with superoxide to form peroxynitrite ($ONOO^-$), which oxidizes and nitrates lipids and proteins. In addition to enzymatic synthesis, nitrite may be converted to NO in the stomach and in the ischemic heart under acidic conditions [10, 11], suggesting an alternative route to nitrosative stress. Additionally, airways are exposed to nitrogen oxides in the external environment [12-14], which impose greater nitrosative stress to the lung in particular in the event of pulmonary injury and asthma [15, 16].

Mammalian cells have developed a multitude of mechanisms to cope with nitrosative stress. DNA damage caused by base deamination may be repaired by a variety of pathways. Several DNA glycosylases such as UNG, SMUG1, TDG, MBD4 and AAG

may be involved in repair of pyrimidine and purine deamination [17-19]. In addition, endonuclease V can remove inosine derived from adenine deamination [20-23]. The nucleotide excision repair pathway may be involved in the repair of dG-dG crosslinks formed by RNS attack. Methionine residues in proteins are susceptible to oxidation by peroxynitrite; however, methionine sulfoxide reductase (msr) is capable of reducing methionine sulfoxide back to the original form [24, 25]. Therefore, a lack of the *msr* gene in *E. coli* confers sensitivity to nitrosative stress [26]. Heme oxygenase-1 (HO-1) is involved in degradation of heme to carbon monoxide, iron and bilirubin. When treated with NO, expression of HO-1 is activated in motor neurons and glial cells, which confers greater resistance to nitrosative stress [27, 28]. This adaptive resistance to NO may be related to the release of bilirubin, which acts as an antioxidant to cleanse NO [29].

To investigate other NO resistance mechanisms in mammalian systems, we developed an *E. coli*-based genetic system to screen for genes involved in resistance to RNS. This system takes advantage of the hypersensitivity of an *E. coli* triple mutant strain (*nfi alkA nei*) to nitrosative stress imposed by RNS. Similar to other previous studies [30, 31], acidified nitrite was used to induce nitrosative stress. In this study, we screened a mouse lung cDNA library for candidate genes involved in resistance to nitrosative stress. Thioredoxin domain-containing 5 (mTrx 5) complemented both the hypersensitivity of the *E. coli* mutant strain to nitrosative stress and the thioredoxin deficiencies in *E. coli*.

III. Materials and Methods

A. Reagents, media and strains

All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO), Fisher Scientific (Suwanee, GA), or VWR (Suwanee, GA). Restriction enzymes, *Taq* DNA polymerase and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). Bovine serum albumin (BSA) and dNTPs were purchased from Promega (Madison, WI). HiTrap chelating columns were purchased from GE Healthcare (Piscataway, NJ). Oligodeoxyribonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA). Sodium nitrite and sodium acetate trihydrate were purchased from Fisher Scientific (Suwanee, GA). LB and SOC media were prepared according to standard recipes [32]. Sonication buffer consisted of 20 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), 0.1 mM DTT (dithiothreitol), 0.15 mM PMSF (phenylmethylsulfonyl fluoride), and 50 mM NaCl. *E. coli* strains BW1466 (CC106) and BW1739 (BW1466 but Δ*nfi*::(FRT-Spc-FRT) *nei*-1::cat *alkA*1) are kind gifts from Dr. Bernard Weiss at Emory University. BW25113, JW5856-2 (BW25113 but Δ *trxA*732::kan) and JW2566-1 (BW25113 but Δ*trxC*750::kan) were obtained from *E. coli* Genetic Stock Center at Yale University.

E. coli host strain BH214 [*thr-1, ara-14, leuB6, tonA31, lacY1, tsx-78, galK2, galE2, dcm-6, hisG4, rpsL, xyl-5, mtl-1, thi-1, ung-1, tyrA*::Tn*10, mug*::Tn*10, supE44*, (DE3)] is a kind gift of Dr. Ashok Bhagwat at Wayne State University. The *E. coli trxA trxC* double deletion strain was derived from JW2566-1 as described below.

B. Construction of mouse lung cDNA library

The Uni-Zap XR mouse lung premade cDNA library was obtained from Stratagene (La Jolla, CA). The Uni-Zap XR vector system accommodates DNA inserts from 0 to 10 kb in length and is designed for the *in vivo* excision of the pBluescript phagemid, allowing the insert to be characterized in a plasmid system. Libraries consisting of >1 x 10^6 cDNA containing phagemid vectors were isolated. The phagemid cDNA library was made by co-infecting XL1-Blue MRF' cells with both the lambda phage (containing the amplified premade library constructed in the Uni-Zap XR vector) and the ExAssist helper phage. The coinfection was performed using 1 x 10^7 pfu of lambda phage (containing library) with a ratio of 1:10:100 for lambda phage: XL1-Blue MRF' cells: helper phage, respectively, in order to ensure complete coverage of the mouse lung cDNA library. Incubation of the co-infected cells for 15 min at 37ºC allowed the phage to attach to the cells, which were then incubated for 2.5 hr at 37ºC with shaking. The cells were lysed by incubation at 70ºC for 20 min and the phage particles isolated by centrifugation to remove the cell debris. To transfer the phagemid particle from the helper phage to an *E. coli* strain, SOLR cells were infected with the library containing helper phage lysate. Since the helper phage contains an amber mutation that prevents replication of the phage genome in a non-suppressing *E. coli* strain (such as SOLR), only the excised phagemid may replicate in the host cell. Infection of cells was performed by addition of 10 to 100 µl (dependent on titer of phage) of the phage lysate to 200 µl of SOLR cells with an O.D. of 1.0 suspended in 10 mM magnesium sulfate, and incubated as described above. The infected cells were then plated onto LB-ampicillin

media and incubated at 37ºC overnight. Colonies were scraped from the plates and plasmid was isolated from the cells to obtain the phagemid cDNA library.

C. Electroporation of BW1739 cells

BW1739 cells were electroporated using an Electro Square Porator ECM 830 (BTX). Electrocompetent cells were made as described previously [32]. A 40-µl aliquot of cells was mixed with 1 µl of the cDNA phagemid library and incubated on ice for 3 min. The mixture was then placed in a sterile 100-µl 1 mm gap electroporation cuvette (USA Scientific) and electroporated with three eight-millisecond pulses at 500 V. The cells were then immediately added to 1 ml of sterile room temperature SOC media and incubated at 37°C for 1 hr with shaking at 200 rpm.

D. Screening of mouse lung library under nitrosative stress

BW1739 cell cultures that were transformed with a portion of mouse lung cDNA plasmid library were plated onto ampicillin containing LB plates to ensure the growth of only plasmid library containing cells. Cell density was approximated using a spectrophotometer and divided into five aliquots of 0.5 ml (containing approximately 3 x 10^8 cell number). Each aliquot was treated with 0, 20, 30, 40 and 50 mM sodium nitrite acidified with acetic acid (pH 4.6) and then incubated at 37°C for 10 min. The treated cells from the 0 mM reaction were diluted 5 x 10^6 times for plating, while treated cell suspensions from the 20 mM to 50 mM reaction were not diluted for plating. BW 1739 cells were transformed with pBluescript SK(+) (an empty vector as a control). Cells were grown to saturation, counted and divided into each tube. The cells were plated onto LB plates with ampicillin (100 µg/ml) and isopropyl-β-D-thiogalactopyranoside (IPTG) (0.5

mM) and the plates were incubated at 37°C for 24 hr. Survival colonies from high concentrations of acidified sodium nitrite treatment were grown in LB-Amp media and plasmids were isolated from each culture. Sequences of the candidate genes in the plasmids were analyzed by DNA sequencing and the sequences obtained were run against the protein sequences database or nucleotide sequence database from GenBank using the FASTX algorithm or BLASTN program.

E. Complementation assay under nitrosative stress

Mouse thioredoxin domain-containing 5 gene (Clone ID # 4911455, Open Biosystems) was digested with SalI and XhoI and cloned in pBluescript to generate pBS-mTrx5. The resulting mTrx 5-containing plasmid was transformed in BW1466, BW1739, BW25113, JW2566-1 (*trxC*$^-$), JW5856-2 (*trxA*$^-$) and the *trxA*$^-$ *trxC*$^-$ deletion strain. The acidified nitrite treatment and plating were conducted as described above in screening of the mouse lung library under nitrosative stress.

F. Construction of E. coli trxA trxC deletion strain

E. coli strain JW2566-1 (BW25113 but Δ*trxC*750::kan) from the Keio collection was used as the recipient cell. To remove the kanamycin cassettee from the recipient cells, plasmid pCP20 was transformed into the JW2566-1 strain. The ampicillin resistant pCP20 contains thermal inducible FLP recombinase and is curable at elevated temperatures due to temperature-sensitive replication [33]. Cells transformed with pCP20 were incubated on LB plates containing 50 μg/ml Amp at 30°C overnight. A few colonies were selected and grown at 42°C on LB plates without antibiotics, with Amp or with Kan to select for the colonies that lost all antibiotic resistance. To construct the

double deletion strain, *E. coli* strain JW5856-2 (BW25113 but Δ*trxA*732::kan) from the Keio collection was used as the donor cell. To make P1 lysate, 2 ml of overnight culture of the donor cells were mixed with 10 µl of P1 *dam rev6* phage suspension ($\geq 10^9$/ml) and incubated for 3 hr at 37°C. For transduction, the donor phage suspension (P1 x JW5856-2) was mixed with 2 ml of the overnight culture of the recipient cells and incubated at 37°C for 20 min to allow for absorption of phage onto the recipient cells. Ten µl of transduction mixture was mixed with 2 ml of LB and incubated at 37°C for 1.5 hr for induction of gene expression and then spread on a LB-Kan plate. The absence of *trxA trxC* genes was confirmed by PCR.

G. Cloning, expression and purification of mouse thioredoxin domain-containing 5

Mouse thioredoxin domain-containing 5 (mTrx 5) gene was amplified by PCR using the forward primer M. Trx 5 F (5'-TGG GGT ACC CAT ATG CCC CCG CGC CCA GGA CGC-3'; the NdeI site is underlined) and the reverse primer M. Trx 5 R (5'-CCC CTC GAG CTA TAG TTC ATC CTT TGC CTG GCG-3'; the XhoI site is underlined). The PCR reaction mixture (50 µl) consisted of 10 ng of mouse thioredoxin domain-containing 5 cDNA (Open Biosystems), 200 nM forward primer M. Trx 5 F and reverse primer M. Trx 5 R, 1 x Taq PCR buffer (New England Biolabs), 200 µM each dNTP, and 5 units of Taq DNA polymerase (New England Biolabs). The PCR procedure included a pre-denaturation step at 94°C for 3 min; 30 cycles of three-step amplification with each cycle consisting of denaturation at 94°C for 50 sec, annealing at 60°C for 50 sec and extension at 72°C for 90 sec; and a final extension step at 72°C for 10 min. The

PCR product was purified with Gene Clean 2 Kit (Qbiogene). The PCR product and the plasmid vector pET16b were digested with NdeI and XhoI, purified with Gene Clean 2 Kit and ligated by T4 DNA ligase. The ligation mixture was transformed into *E. coli* strain JM109 competent cells by electroporation method [32]. The sequence of the mTrx 5 gene in the resulting plasmid (pET16b-mTrx5) was confirmed by DNA sequencing.

To express the N-terminal His-6-tagged mTrx 5 protein, pET16b-mTrx5 plasmid was transformed into *E. coli* strain BH214 (DE3) by electroporation [32]. An overnight *E. coli* culture was diluted 100-fold into an LB medium (1 L) supplemented with 100 μg/ml ampicillin. The *E. coli* cells were grown at 37°C with shaking at 250 rpm until the optical density at 600 nm reached approximately 0.6. IPTG was added to a final concentration of 1 mM. The culture was allowed to grow at room temperature for an additional 16 hr. The cells were collected by centrifugation at 4,000 rpm with JS-4.2 rotor in J6-MC centrifuge (Beckman Coulter) at 4°C and washed once with pre-cooled sonication buffer.

To purify the mTrx 5 protein, the cell pellet from the one liter culture was suspended in 10 ml of sonication buffer and sonicated at output 5 for 3 x 1 min with 5 min rest on ice between intervals using a Sonifier Cell Disruptor 350 (Branson). The sonicated solution was clarified by centrifuging at 12,000 rpm in a SS-34 rotor at 4°C for 20 min. Since the mTrx 5 protein was found in the inclusion bodies, the pellet was denatured by urea, purified by Ni-NTA agarose beads and refolded by dialysis [34]. Briefly, the pellet from sonication was re-suspended in 5 ml of Lysis Buffer B (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 8.0) and stirred at room temperature for 1 hr.

The solution was centrifuged by 10,000 rpm at 4°C for 20 min to remove precipitates. One ml of the 50% Ni-NTA agarose (Qiagen) was added in the supernatant and shaken in a Quake shaker (Barnstead) for 1 hr at room temperature. The lysate–resin mixture was carefully loaded into a 3-ml empty column, washed twice with 3 ml of Buffer C (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 6.3), and three times with 1 ml of Buffer D (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 5.9). The proteins were eluted four times with 0.5 ml of Buffer E (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 4.5) and collected in four 0.5 ml fractions. Each fraction from Buffer C, D and E was examined by 10% SDS-PAGE. The fractions containing mTrx 5 protein (fractions 1-3 from Buffer E elution) were pooled, dialyzed in dialysis buffer (20 mM Tris-HCl, 1 mM EDTA, 0.1 mM DTT, pH 8.0) at 4°C overnight, and concentrated using an YM-10 microcon column (Millipore). The amount of protein was quantified on a 10% SDS-PAGE using BSA as a standard.

H. Cloning, expression and purification of *E. coli* thioredoxins 1 and 2

Since pET32a (Novagen) contains *E. coli trxA* gene (encodes thioredoxin 1), it was directly used for expression of thioredoxin 1 protein. *E. coli trxC* gene (encodes thioredoxin 2) was amplified by PCR using the forward primer E.Trx2 F (5'-TGG GGT ACC CAT ATG AAT ACC GTT TGT ACC CAT - 3'; the NdeI site is underlined) and the reverse primer E. Trx2 R (5'-CCC GGA TCC TTA AAG AGA TTC GTT CAG CCA-3'; the BamHI site is underlined). The PCR reaction mixture (50 µl) consisted of 8 ng of *E. coli* genomic DNA, 200 nM forward primer E.Trx2 F and reverse primer E.Trx2 R, 1 x Taq PCR buffer (New England Biolabs), 200 µM each dNTP, and 5 units of Taq

DNA polymerase (New England Biolabs). The PCR procedure included a pre-denaturation step at 94°C for 5 min; 30 cycles of three-step amplification with each cycle consisting of denaturation at 94°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 1.5 min; and a final extension step at 72°C for 10 min. The remaining molecular cloning procedures were the same as described above for mTrx 5. The sequence of the *E. coli trxC* gene in the resulting plasmid (pET16b-E.Trx2) was confirmed by DNA sequencing.

The protein expression was induced by 1 mM IPTG as described above for mTrx 5. To purify the *E. coli* thioredoxin 2 protein, the cell pellet from a 500 ml culture was suspended in 10 ml of sonication buffer and sonicated at output 5 for 3 x 1 min with 5 min rest on ice between intervals. The sonicated solution was clarified by centrifuging at 12,000 rpm at 4°C for 20 min. The supernatant was transferred into a fresh tube and loaded into a 1 ml HiTrap chelating column. The bound protein in the column was eluted with a linear gradient of five column volumes of 0-0.5 M imidazole in chelating buffer B (20 mM sodium phosphate (pH 7.4), 500 mM NaCl). Fractions of the eluate were analyzed by 10% SDS-PAGE, and those fractions containing thioredoxin 2 protein (0.2-0.3 M imidazole) were pooled and concentrated using an YM-10 microcon column (Millipore). The amount of protein was quantified on a 10% SDS-PAGE, using BSA as a standard. The *E. coli* thioredoxin 1 protein in the pET32a was induced and purified in the same manner.

I. Plasmid nicking assay

The reaction mixtures (200 µl) containing 3 µg of plasmid DNA pBluescript, 1 mM acidified nitrite (pH 4.6), 10 mM potassium phosphate buffer (pH 4.7), and 4 µM protein (*E. coli* thioredoxin 1, *E. coli* thioredoxin 2, mTrx 5 or carbonic anhydrase) were incubated at room temperature for 16 hr. An equal volume of isopropanol was then added to the reaction mixtures and stored at -80°C for 1 hr to precipitate DNA. The DNA was collected by centrifugation and dissolved in 20 µl of deionized water. The DNA solutions (10 µl) were supplemented with *E. coli* endonuclease V protein to 100 nM final concentration in a buffer containing 10 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$ and incubated at 37°C for 1 hr. The reaction mixtures (13 µl) were mixed with 3 µl of 6 x agarose gel loading buffer containing 30% glycerol, 0.1% bromophenol blue. The different forms of plasmid after the treatment were separated by electrophoresis on a 1 % agarose gel and visualized by ethidium bromide staining.

J. Statistical analysis

One-way and two-way analysis of variance (ANOVA) were performed using Programming language R. The differences are considered statistically significant with *p*-value less than 0.05.

IV. Results

A. E. coli-based genetic system

To investigate the resistance mechanisms to RNS adopted by mammalian cells, we sought to develop a genetic system to screen genes involved in resistance to RNS. We tested the idea of establishing an *E. coli* strain that is hypersensitive to nitrosative stress.

Since endo V (*nfi*), alkA and endo VIII are involved in repair of DNA damage

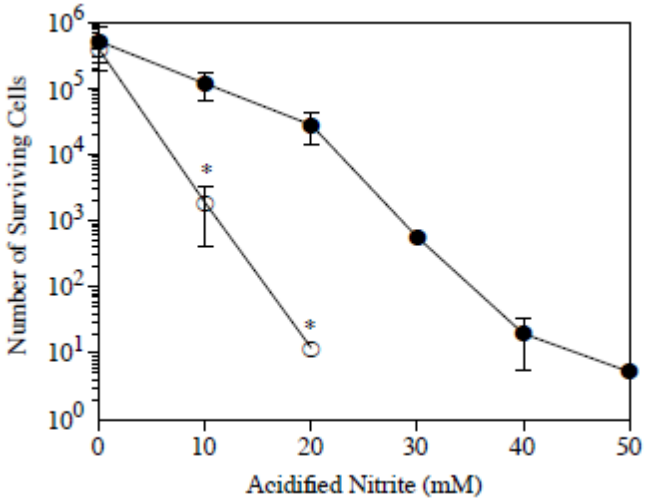

Figure 2.1 Sensitivity of *E. coli* strains BW1466 and BW1739 to acidified sodium nitrite treatment. Colonies from each strain were grown to saturation in 3.5 ml of LB medium. Cell density was approximated using a hemocytometer and divided into six aliquots of 0.5 ml (containing approximately 5 x 10^5 cell number). Each aliquot was treated with 0, 10, 20, 30, 40, and 50 mM acidified nitrite (pH 4.6) and then incubated at 37°C for 10 min. Following incubation, small portions (10 µl) of the cell mixtures were diluted into 1 ml of cold LB media to stop the reaction (50 µl and 100 µl for the 40 and 50 mM, respectively). Cells from the 0 and 10 mM reactions were diluted 2 additional times, 20 mM once more and the remaining reactions were not further diluted. Each resulting mixture was plated by adding 50 µl of the diluted mixture onto LB plates. Plates were incubated overnight at 37ºC and colonies counted. (●) wt *E. coli* strain (BW1466); **(○)** DNA repair deficient *E. coli* strain (BW1739). Data were obtained from five independent experiments and expressed as means ± SD. All the three *p*-values from two-way ANOVA were less than 0.05.

caused by RNS, we surmised that deleting these repair genes might sensitize the cells to nitrosative stress. *E. coli* BW1739 lacks endo V alkA and endo VIII. To test the sensitivity of this triple mutant strain to nitrosative stress, we treated it with different concentrations of acidified nitrite. Consistent with our hypothesis, the triple mutant strain BW1739 was much more sensitive than the corresponding wt (wild type) strain BW1466 (Fig. 2.1). After treating with more than 20 mM acidified nitrite, BW1739 completely

lost viability (Fig. 2.2A), indicating that the triple mutant strain BW1739 became hypersensitive to nitrosative stress.

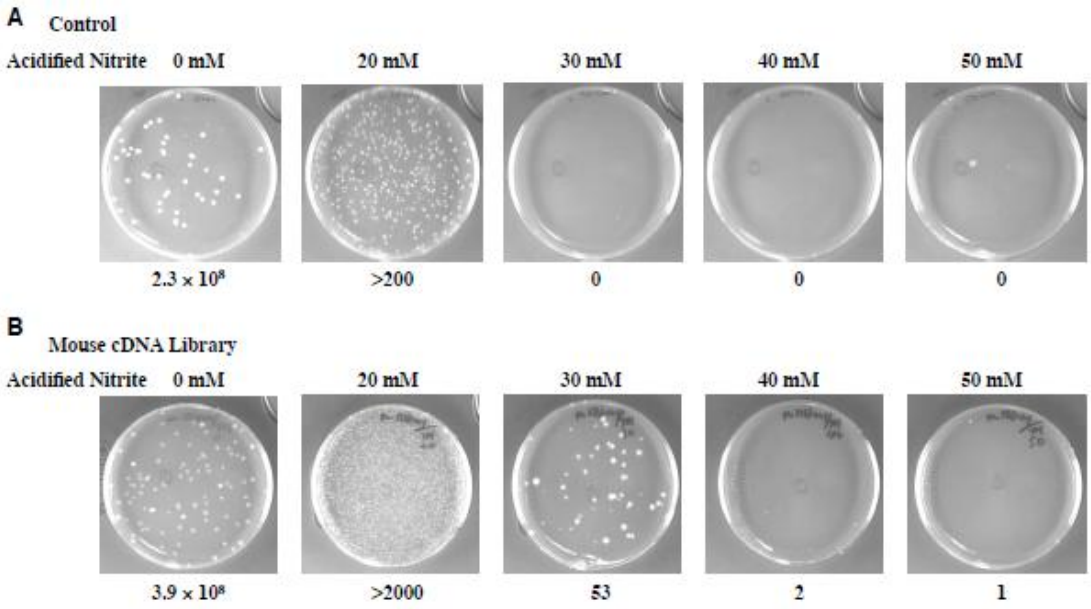

Figure 2.2 Screening of mouse lung cDNA library with acidified sodium nitrite. See Materials and Methods for details. A. E. coli strain BW1739 containing empty pBluescript vector. B. E. coli strain BW1739 containing mouse lung cDNA library.

B. Screening of candidate genes in mouse lung cDNA library

We reasoned that genes involved in resistance to RNS might complement the hypersensitive phenotype demonstrated in the triple mutant strain. A pre-made commercial mouse lung cDNA phage library was converted to a plasmid expression library as detailed in Materials and Methods. Mouse lung cDNA was chosen because the lung is constantly exposed to environmental nitrosative stress. After electroporation into the BW1739 mutant strain, cells were treated with various concentrations of acidified sodium nitrite and plated out on LB-Amp plates. The acidic condition or the 50 mM sodium nitrite alone did not cause significant reduction in survival (Fig. 2.2, 0 mM and

Fig. S1). While cells with empty vector as control did not show survival above 20 mM concentrations, some colonies in the mouse cDNA library did (Fig. 2.2B). Plasmids from those surviving colonies were then isolated and subject to DNA sequencing to identify the candidate genes. The DNA sequences matched thioredoxin domain-containing 5 (Genebank accession no. NP_663342), Zcchc14 protein (AAH05674), lysozyme (NP_059068), macrophage migration inhibitory factor (AAH86928.1), Zfp265 (AAF04474), hypothetical protein LOC319587 (NP_795929.1), ubiquitination factor E4A (NP_663375.2), decorin (NP-031859.1), armadillo repeat-containing protein (NP_083116.1), unnamed protein product (CAA76142.1), unnamed protein product (BAC29280.1), and unnamed protein product (BAC28255.1). The inability to identify

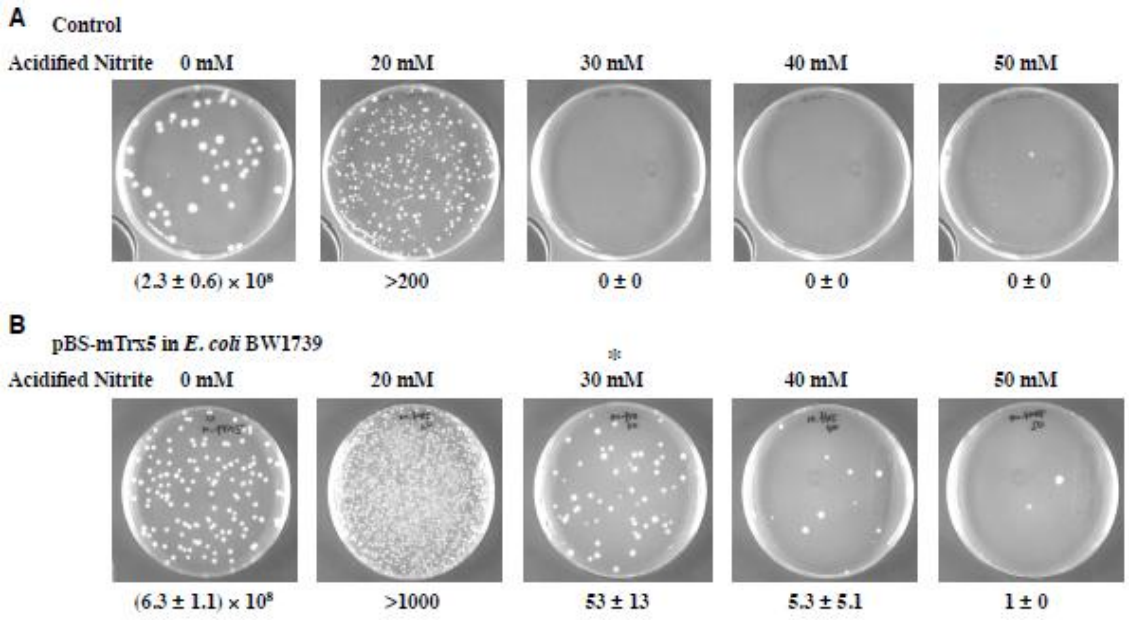

Figure 2.3 The resistance of mTrx 5 to nitrosative stress. See Materials and Methods for details. A. E. coli strain BW1739 containing empty pBluescript vector. B. E. coli strain BW1739 containing pBS-mTrx 5 vector. Data shown are a representative of three independent experiments and expressed as means ± SD when possible. Statistical analysis was performed using one-way ANOVA. *: $p < 0.05$.

DNA repair genes in the screening may be caused by incomplete screening and/or a growth inhibition effect exhibited by DNA repair enzymes under a nitrosative stress background ([35] and data not shown).

C. Verification of mouse thioredoxin domain-containing 5

One of the genes that caught our particular attention was the mouse thioredoxin domain-containing 5 (mTrx 5), which contains three thioredoxin domains in a span of 417 amino acids. To verify the resistance of mTrx 5 to nitrosative stress, the mTrx 5 gene was cloned into pBluescript and transformed to the BW1739 strain. After treating with acidified nitrite, the mTrx 5-containing cells were able to survive much better than the control (Fig. 2.3). While the control showed a much reduced number of colonies at 20 mM concentration and no survival above 20 mM, mTrx 5-containing cells were much more viable at the high level of nitrosative stress. Since the mTrx 5 gene was retransformed into the BW1739 strain, this indicated that the survival in the initial screen was not due to chromosomal changes in the BW1739 host.

D. Reduction of RNS-induced DNA damage by *E. coli* thioredoxins and mTrx 5

Since RNS inflict DNA damage and the BW1739 strain is deficient in deaminated DNA repair, we thought one of the possible mechanisms of resistance to nitrosative stress was to reduce RNS-induced DNA damage. To test this possibility, we cloned, expressed, and purified mTrx 5 protein along with *E. coli* Trx 1 and Trx 2 proteins (Fig. S2). A plasmid-based nicking assay was used to test the effect of these proteins on nicking of plasmid treated with acidified nitrite (Fig. 2.4A). For each pair, the treatment of the plasmid was carried out in the presence of Trx 1, Trx 2 or mTrx 5. After the treatment,

one tube from each pair was treated with *E. coli* endo V to assess DNA damage, because endo V is capable of nicking all deaminated base lesions [21, 23, 36-40].

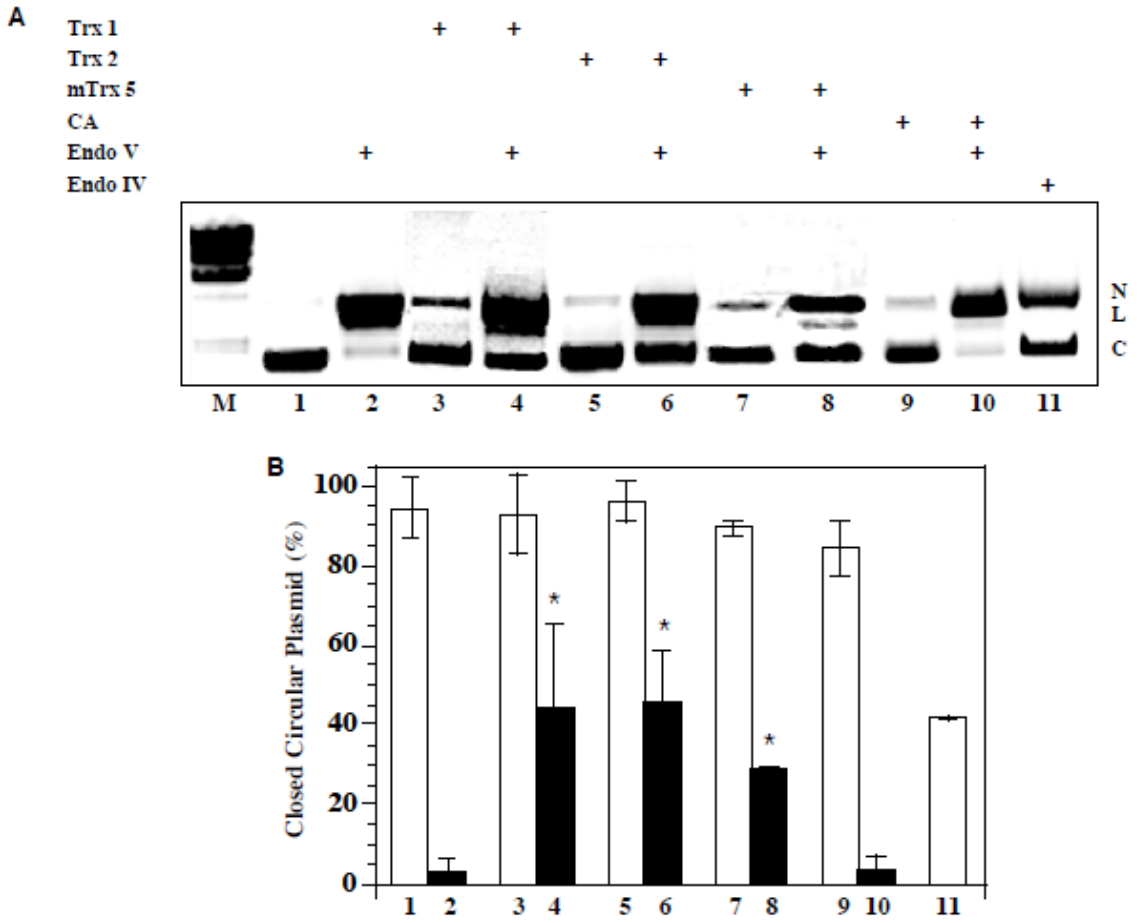

Figure 2.4 Plasmid nicking assay with E. coli thioredoxins 1 and 2 and mTrx 5 after acidified nitrite treatment. A. Agarose gel electrophoresis analysis of plasmid nicking products. pBluescript was treated with acidified nitrite in the presence of E. coli thioredoxin 1, thioredoxin 2 or mTrx 5 as described in Materials and Methods. M: lambda DNA Hind III digest. CA: carbonic anhydrase. Lane 1 and 2: no other protein added (except endo V to lane 2); lanes 3 and 4: E. coli thioredoxin 1 protein added; lanes 5 and 6: E. coli thioredoxin 2 protein added; lanes 7 and 8: mTrx 5 protein added; lanes 9 and 10: carbonic anhydrase (Sigma, C-7025). After acidified nitrite treatment, samples from lanes 2, 4, 6, 8 and 10 were incubated with E. coli endo V at 37°C for 1 hr. Lane 11: identical to the sample in lane 2 but treated with E. coli endonuclease IV instead of endo V. After acidified nitrite treatment and isopropanol precipitation, the sample (10 µl) was incubated with 100 nM E. coli endo IV in a buffer containing 10 mM Tris-HCl (7.0), 10 mM MgCl2, 1 mM DTT at 37°C for 1 hr. C: closed circular plasmid; N: nicked plasmid; L: linear plasmid. B. Plot of remaining closed circular plasmid. Data were obtained from two independent experiments and expressed as means ± SD. Statistical analysis was performed using one-way ANOVA. *: $p < 0.05$.
See Materials and Methods for details. A. E. coli strain BW1739 containing empty pBluescript vector. B. E. coli strain BW1739 containing pBS-mTrx 5 vector. Data shown are a representative of three independent experiments and expressed as means ± SD when possible. Statistical analysis was performed using one-way ANOVA. *: $p < 0.05$.

In the control reaction without any other protein, treatment with acidified nitrite apparently generated many endoV recognizable lesions in the plasmid (Fig. 2.4A, lanes 1-2). Consequently, after endo V treatment, most closed circular plasmids were converted to nicked and linear forms with approximately 3% left as the closed circular form (Fig. 2.4B). When the treatment was performed in the presence of *E. coli* Trx 1 protein, the amount of remaining closed circular form increased to approximately 30% (Fig. 2.4B, lanes 3-4). Likewise, addition of *E. coli* Trx 2 protein to the reaction increased the closed circular form to approximately 45% (Fig. 2.4B, lanes 5-6). When mTrx 5 protein was present in the treatment mixture, the remaining closed circular form increased to approximately 30% (Fig. 2.4B, lanes 7-8). As another control, we also performed the same assay with carbonic anhydrase. In this case, no apparent reduction in DNA damage was observed (Fig. 2.4B, lanes 9-10). Because endo V has AP endonuclease activity [21, 22], we used the known AP endonuclease endo IV to assess the percentage of plasmid nicking in lane 2 that is related to formation of AP sites. More than 50% of the acidified nitrite treated plasmid was converted to nicked plasmid (Fig. 2.4, lane 11), suggesting significant generation of AP sites. This is understandable given that the treatment was carried out under acidic conditions, which is known to accelerate depurination. These results indicated that more than 40% of the RNS-mediated damage is non-AP site damage, and the majority of which is likely due to deamination.

E. Complementation of *trxA trxC* by mTrx 5

Since both the thioredoxin 1 and thioredoxin 2 proteins were able to reduce the RNS-induced DNA damage, we set out to test whether the lack of these genes could

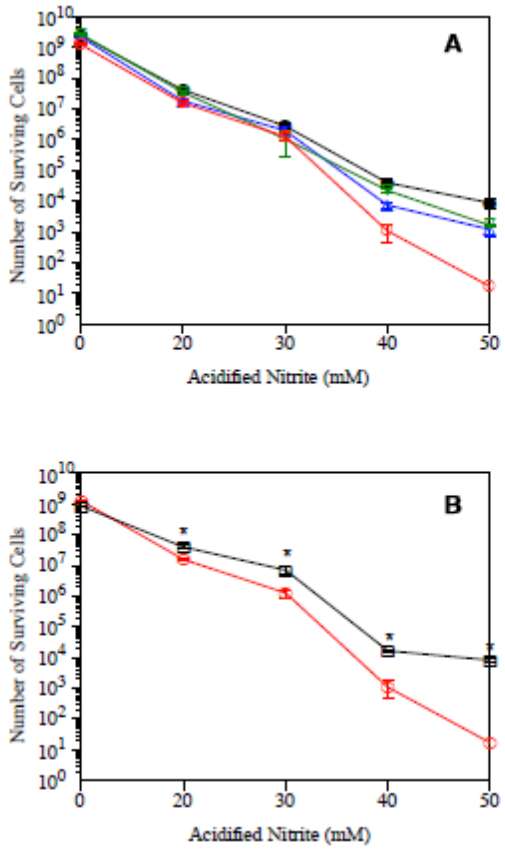

Figure 2.5 Complementation of trxA and trxC by mTrx 5 under nitrosative stress. The treatment with acidified nitrite and plating were performed as detailed in Materials and Methods. A. Sensitivity of E. coli trxA and trxC to nitrosative stress. (□, black) Wild type E. coli strain (BW25113) containing pBluescript vector; (□, blue) E. coli trxA deletion strain (JW5856-2) containing pBluescript vector; (□, green) E. coli trxC deletion strain (JW2566-1) containing pBluescript vector; (□, red) E. coli trxA trxC double deletion strain (BW25113 but trxA- trxC-) containing pBluescript vector. Data were obtained from three independent experiments and expressed as means ± SD. Two-way ANOVA was performed between wt and trxA deletion ($p < 0.05$), wt and trxC deletion ($p < 0.05$), and wt and trxA trxC double deletion ($p < 0.05$). B. Complementation of mTrx 5 to E. coli trxA trxC double deletion strain. (□, red) E. coli trxA trxC double deletion strain (BW25113 but trxA- trxC-) containing pBluescript vector; (□, black) E. coli trxA trxC double deletion strain (BW25113 but trxA- trxC-) containing pBS-mTrx5 vector. Data were obtained from three independent experiments and expressed as means ± SD. Statistical analysis was performed using two-way and one-way ANOVA. *: $p < 0.05$. Post tests after two-way ANOVA were performed using the Bonferroni method. The differences between each concentration and 0 mM were significant ($p < 0.05$) for both the E. coli trxA trxC double deletion strain and the same strain containing pBS-mTrx5 vector.

result in sensitivity to nitrosative stress and whether mTrx 5 could complement them. Lack of *trxA* (encodes thioredoxin 1) or *trxC* (encodes thioredoxin 2) appeared to cause a small reduction in resistance to nitrosative stress at high acidified nitrite concentrations (Fig. 2.5A). Deletion of both *trxA* and *trxC* genes led to a significantly heightened sensitivity to nitrosative stress (Fig. 2.5A). To test whether mTrx 5 can complement the *trxA trxC* phenotype *in vivo*, we cloned the mTrx 5 gene to pBluescript and transformed it into the double deletion strain. The presence of the mTrx 5 gene appeared to revert the *trxA trxC* phenotype by rendering it significantly more resistant to nitrosative stress (Fig. 2.5B). These results suggest that mTrx 5 could substitute some of the biological functions of *trxA* and *trxC* in the *E. coli* host.

F. Reduction of ROS-induced DNA damage by *E. coli* thioredoxins and mTrx 5

Since thioredoxin is known to scavenge RNS, we tested the ability of E. coli thioredoxins 1 and 2, and mTrx 5 to reduce ROS-induced DNA damage in the plasmid-based nicking assay. Prior to ROS treatment, the majority of the plasmid was in the closed circular form (Fig. 2.6A, lane C). After the treatment, the majority of the plasmid was converted to nicked/linear forms, indicating significant strand-breaking damage caused by ROS (Fig. 2.6A, lane 2). The remaining closed circular plasmid apparently harbored Fpg-sensitive damage, which was converted to nicked/linear forms by the bifunctional DNA glycosylase Fpg (Fig. 2.6A, lane 3). Addition of E. coli thioredoxins 1 and 2, and mTrx 5 all significantly reduced the strand breaks in the absence of Fpg (Fig. 2.6A, compare lanes 3, 5 and 7 with lane 1; Fig. 2.6B). Furthermore, thioredoxins 1 and 2, and mTrx 5 significantly reduced formation of Fpg-sensitive sites (Fig. 2.6A, compare

lanes 4, 6 and 8 with lane 2; Fig. 2.6B). These results suggest that mTrx 5, like E. coli thioredoxins 1 and 2, can protect DNA from ROS-induced damage.

V. Discussion

Cells are constantly exposed to reactive nitrogen species from the environment and endogenous sources and, under such conditions, must adopt various mechanisms to cope with nitrosative stress. One of the antinitrosative stress mechanisms is to repair DNA damage generated by RNS. Indeed, macrophage-residing bacterial pathogens seem to utilize this mechanism well. In *Salmonella typhimurium*, RecB and RecC of the recombinational repair pathway affect its survival in macrophages [41]. In *Mycobacterium tuberculosis* and S. *typhimurium*, lack of uvrB of the nucleotide excision repair pathway contributes to mutations and sensitivity caused by nitrosative stress [30, 42-44]. An *E. coli* strain such as BW1739 lacking DNA repair glycosylases is also hypersensitive to RNS (Fig. 2.1), underscoring the importance of DNA repair in attenuating the genotoxicity of RNS-induced DNA damage. Taking advantage of the hypersensitive phenotype, this study screened a mouse lung cDNA library to identify mammalian genes that may confer resistance to nitrosative stress. Among the genes identified from the library, this work focused on mouse thioredoxin domain-containing 5, a gene which contains three thioredoxin-like domains. Thioredoxin was discovered in *E. coli* as a small redox protein back in 1964 [45]. Since then, thioredoxin proteins have been found in a variety of bactacterial, archaeal and eukaryotic organisms [46-48]. *E. coli* genome contains a second thioredoxin which plays a similar role as thioredoxin 1 but is induced by H_2O_2 in an OxyR-dependent manner [49-51].

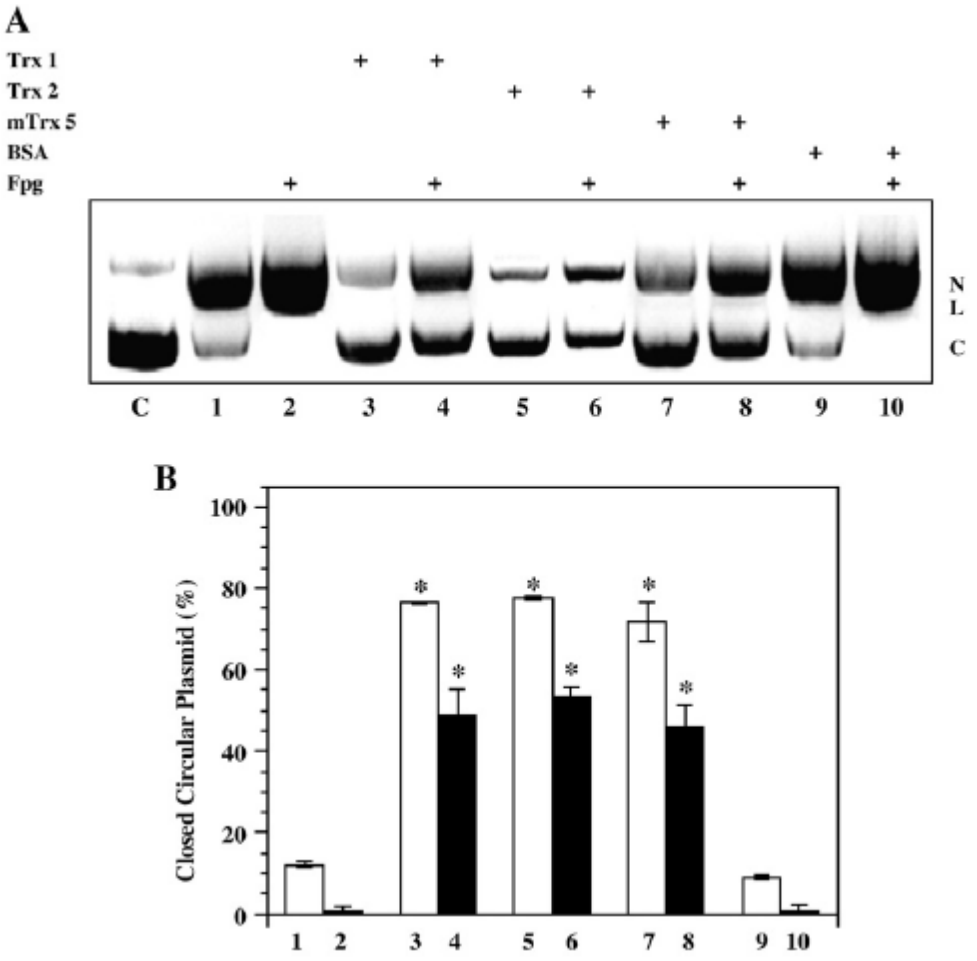

Figure 2.6 Plasmid nicking assay with *E. coli* thioredoxins 1 and 2 and mTrx 5 after $H_2O_2/FeCl_3$ treatment. A. Agarose gel electrophoresis analysis of plasmid nicking products. The assay was carried out similar to a previous procedure [65]. Reaction mixtures (50 µl) containing 12.5 mM Tris-HCl (pH 8.0), 12.5 mM NaCl, 0.1 mM $FeCl_3$, 0.4 mM EDTA, 10 mM H_2O_2, 3 µg of pBluescript, and 4 µM purified protein were incubated at room temperature for 1 hr. The reactions were stopped by addition of dipyridyl (iron chelator) to a final concentration of 5 mM. An equal volume of isopropanol was added to the reaction mixtures and stored at -80°C for 1 hr to precipitate DNA. After centrifugation, the pellets were dissolved in 50 µl of deionized water. The plasmid samples (10 µl) were then incubated with 100 nM *E. coli* Fpg (MutM) protein (NEB) in a buffer containing 100 mM KCl, 5 mM EDTA, 2 mM 2-mercaptoethanol at 37°C for 1 hr. The reaction mixtures (13 µl) were mixed with 3 µl of 6 x agarose gel loading buffer, electrophoresed in 1% agarose gels and visualized by ethidium bromide staining. C: control treatment without H_2O_2 and Fpg. Lanes 1 and 2: no protein added; lanes 3 and 4: *E. coli* thioredoxin 1 protein added; lanes 5 and 6: E.coli thioredoxin 2 protein added; lanes 7 and 8: m.Trx5 protein added; lanes 9 and 10: bovine serum albumin (Sigma A-8531). After $H_2O_2/FeCl_3$ treatment, samples from lanes 2, 4, 6, 8, and 10 were incubated with *E. coli* Fpg as described above. C: closed circular plasmid; N: nicked plasmid; L: linear plasmid. B. Plot of remaining closed circular plasmid. Data were obtained from two independent experiments and expressed as means ± SD. Pair-wise comparisons (lanes 3 and 1, lanes 5 and 1, lanes 7 and 1, lanes 9 and 1; lanes 4 and 2, lanes 6 and 2, lanes 8 and 2, and lanes 10 and 2) were performed using one-way ANOVA. *: $p < 0.05$.

Mammalian genomes encode multiple thioredoxin and thioredoxin-like proteins [52, 53]. Many of these proteins are involved in response to oxidative stress, protein folding, and acting as a hydrogen donor or disulfide reductase. The function of mTrx 5 *in vivo* and *in vitro* is not well known. Complementation analysis in yeast suggests that mTrx 5 plays the role of a protein disulfide isomerase [54]. The finding of mTrx 5 in our genetic screening for resistance to nitrosative stress is in general consistent with the multiple functions a thioredoxin-like protein may perform.

The relationship between thioredoxin and oxidative stress is extensively studied [46-48, 55, 56]. Reactive oxygen species can be scavenged either directly by thioredoxin and thioredoxin-related protein TRP14 or through reduction of peroxiredoxin [57-60]. In addition, thioredoxin can regenerate glyceraldehyde-3-phosphate dehydrogenase inactivated by H_2O_2 treatment [56, 57], suggesting that thioredoxin can also act as a protein repair enzyme.

The relationship between thioredoxin and thioredoxin-like proteins and nitrosative stress is not well understood. This work provides experimental evidence that *E. coli* thioredoxin and thioredoxin-like proteins such as mTrx 5 are involved in resistance to nitrosative stress. The expression level of mTrx5 in mice is not clear. According to a previous report, it appears that the expression of thioredoxin domain-containing 5 protein is relatively higher in lung tissues than other tissues [61], which may provide protection to nitrosative stress in lung. Because mTrx 5 was obtained through screening of a DNA repair deficient strain, we thought one of the possible mechanisms is that mTrx can reduce the DNA damage caused by exposure to RNS. Indeed, in a plasmid nicking assay,

mTrx 5 protein reduced the nicking of endo V-sensitive damage (Fig. 2.4). Interestingly, similar observations were made of *E. coli* thioredoxins 1 and 2. A simple explanation of these data is that thioredoxins and thioredoxin-like proteins may be able to scavenge the RNS generated from acidified nitrite. If so, thioredoxins and thioredoxin-like proteins are able to scavenge both reactive oxygen species and reactive nitrogen species; in doing so, the RNS-mediated DNA damage is reduced. By scavenging RNS, thioredoxins and thioredoxin-like proteins may also alleviate cytotoxicity by reducing damage to other macromolecules such as proteins.

The contribution of *E. coli* thioredoxins 1 and 2 to resistance toward nitrosative stress was tested *in vivo* using deletion strains. A single deletion of *trxA* or *trxC* conferred modest sensitivity to nitrosative stress, whereas the double deletion *trxA trxC* strain become much more sensitive (Fig. 2.5A). These results are consistent with a study of bacterial pathogen *Helicobacter pylori*, in which the *trx1 trx2* double deletion strain was hypersensitive to nitrosative stress [62]. However, trx1 in *H. pylori* seems to contribute in large part to the sensitivity. The defect caused by *trxA trxC* double deletion was compensated for by the mTrx 5 (Fig. 2.5B), suggesting that mTrx 5 can substitute for the function of *trxA* and *trxC* under nitrosative stress. While this may be due to the ability of mTrx 5 to reduce endo V-sensitive DNA damage, it is possible that the antinitrosative function of mTrx 5 goes beyond simply scavenging the RNS.

The other possibilities are discussed as follows. First, cellular proteins are subject to S-nitrosylation under nitrosative stress [63]. The S-nitrosylation has a profound impact on the physiological function of many proteins. Thioredoxin is one of the

physiological denitrosylases that can remove nitric oxide from a variety of proteins [64]. The denitrosylation process mediated by thioredoxin and possibly by mTrx 5 may help prevent loss of functions in many proteins. Second, under oxidative and nitrosative stress conditions, NO can react with superoxide to generate peroxynitrite, which then oxidizes methionine to methionine sulfoxide. Methionine sulfoxide reductase (Msr) is responsible for reverting methionine sulfoxide back to methionine [24, 25]. During the reaction, Msr is oxidized to form a disulfide bond and requires thioredoxin to reduce it. *E. coli* becomes hypersensitive to nitrosative and oxidative stress without msrA [26]. The lack of thioredoxins in the *trxA trxC* double deletion strain may impede methionine repair and therefore contribute to its sensitivity toward nitrosative stress. It remains to be determined whether thioredoxin-like proteins such as mTrx 5 can participate in the catalytic cycle of methionine damage. The involvement of these mechanisms in resistance to nitrosative stress requires further investigation.

VI. Acknowledgements

This project was supported in part by Cooperative State Research, Education, and Extension Service/United States Department of Agriculture (SC-1700274, technical contribution No. 5768), Department of Defense-Army Research Office (W911NF-05-1-0335 and W911NF-07-1-0141), Calhoun Honors College of Clemson University and Howard Hughes Medical Institute-SC-Life. We thank Dr. Bernard Weiss at Emory University and members of the Cao laboratory for stimulating discussions. We also thank Dr. Brian Dominy for editorial help.

VII. References

[1] Wink, D. A.; Mitchell, J. B. Chemical biology of nitric oxide: Insights into regulatory, cytotoxic, and cytoprotective mechanisms of nitric oxide. *Free Radic Biol Med* **25:**434-456; 1998.

[2] Nathan, C.; Shiloh, M. U. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc Natl Acad Sci U S A* **97:**8841-8848.; 2000.

[3] Griffith, O. W.; Stuehr, D. J. Nitric oxide synthases: properties and catalytic mechanism. *Annu Rev Physiol* **57:**707-736; 1995.

[4] MacMicking, J.; Xie, Q. W.; Nathan, C. Nitric oxide and macrophage function. *Annu Rev Immunol* **15:**323-350; 1997.

[5] Marletta, M. A. Nitric oxide synthase: aspects concerning structure and catalysis. *Cell* **78:**927-930; 1994.

[6] Bogdan, C. Nitric oxide and the immune response. *Nat Immunol* **2:**907-916; 2001.

[7] Fang, F. C. Perspectives series: host/pathogen interactions. Mechanisms of nitric oxide-related antimicrobial activity. *J Clin Invest* **99:**2818-2825; 1997.

[8] Firmani, M. A.; Riley, L. W. Reactive nitrogen intermediates have a bacteriostatic effect on Mycobacterium tuberculosis in vitro. *J Clin Microbiol* **40:**3162-3166; 2002.

[9] Dedon, P. C.; Tannenbaum, S. R. Reactive nitrogen species in the chemical biology of inflammation. *Arch Biochem Biophys* **423:**12-22; 2004.

[10] Weitzberg, E.; Lundberg, J. O. Nonenzymatic nitric oxide production in humans. *Nitric Oxide* **2:**1-7; 1998.

[11] Zweier, J. L.; Wang, P.; Samouilov, A.; Kuppusamy, P. Enzyme-independent formation of nitric oxide in biological tissues. *Nat Med* **1:**804-809; 1995.

[12] Brunekreef, B.; Holgate, S. T. Air pollution and health. *Lancet* **360:**1233-1242; 2002.

[13] Jenkins, H. S.; Devalia, J. L.; Mister, R. L.; Bevan, A. M.; Rusznak, C.; Davies, R. J. The effect of exposure to ozone and nitrogen dioxide on the airway response of atopic asthmatics to inhaled allergen: dose- and time-dependent effects. *Am J Respir Crit Care Med* **160:**33-39; 1999.

[14] van der Vliet, A.; Eiserich, J. P.; Shigenaga, M. K.; Cross, C. E. Reactive nitrogen species and tyrosine nitration in the respiratory tract: epiphenomena or a pathobiologic mechanism of disease? *Am J Respir Crit Care Med* **160:**1-9; 1999.

[15] Allen, B. W.; Demchenko, I. T.; Piantadosi, C. A. Two faces of nitric oxide: implications for cellular mechanisms of oxygen toxicity. *J Appl Physiol* **106:**662-667; 2009.

[16] Andreadis, A. A.; Hazen, S. L.; Comhair, S. A.; Erzurum, S. C. Oxidative and nitrosative events in asthma. *Free Radic Biol Med* **35:**213-225; 2003.

[17] Krokan, H. E.; Drablos, F.; Slupphaug, G. Uracil in DNA--occurrence, consequences and repair. *Oncogene* **21:**8935-8948; 2002.

[18] Wuenschell, G. E.; O'Connor, T. R.; Termini, J. Stability, miscoding potential, and repair of 2'-deoxyxanthosine in DNA: implications for nitric oxide-induced mutagenesis. *Biochemistry* **42:**3608-3616; 2003.

[19] Wyatt, M. D.; Allan, J. M.; Lau, A. Y.; Ellenberger, T. E.; Samson, L. D. 3-methyladenine DNA glycosylases: structure, function, and biological importance. *Bioessays* **21:**668-676.; 1999.

[20] Moe, A.; Ringvoll, J.; Nordstrand, L. M.; Eide, L.; Bjoras, M.; Seeberg, E.; Rognes, T.; Klungland, A. Incision at hypoxanthine residues in DNA by a mammalian homologue of the Escherichia coli antimutator enzyme endonuclease V. *Nucleic Acids Res* **31:**3893-3900; 2003.

[21] Yao, M.; Hatahet, Z.; Melamede, R. J.; Kow, Y. W. Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from Escherichia coli. *J Biol Chem* **269:**16260-16268; 1994.

[22] Huang, J.; Lu, J.; Barany, F.; Cao, W. Multiple Cleavage Activities of Endonuclease V from Thermotoga maritima: Recognition and Strand Nicking Mechanism. *Biochemistry* **40:**8738-8748.; 2001.

[23] Guo, G.; Weiss, B. Endonuclease V (nfi) mutant of Escherichia coli K-12. *J Bacteriol* **180:**46-51; 1998.

[24] Boschi-Muller, S.; Gand, A.; Branlant, G. The methionine sulfoxide reductases: Catalysis and substrate specificities. *Arch Biochem Biophys* **474:**266-273; 2008.

[25] Zhang, X. H.; Weissbach, H. Origin and evolution of the protein-repairing enzymes methionine sulphoxide reductases. *Biol Rev Camb Philos Soc* **83:**249-257; 2008.

[26] St John, G.; Brot, N.; Ruan, J.; Erdjument-Bromage, H.; Tempst, P.; Weissbach, H.; Nathan, C. Peptide methionine sulfoxide reductase from Escherichia coli and Mycobacterium tuberculosis protects bacteria against oxidative damage from reactive nitrogen intermediates. *Proc Natl Acad Sci U S A* **98:**9901-9906; 2001.

[27] Bishop, A.; Marquis, J. C.; Cashman, N. R.; Demple, B. Adaptive resistance to nitric oxide in motor neurons. *Free Radic Biol Med* **26:**978-986; 1999.

[28] Kitamura, Y.; Furukawa, M.; Matsuoka, Y.; Tooyama, I.; Kimura, H.; Nomura, Y.; Taniguchi, T. In vitro and in vivo induction of heme oxygenase-1 in rat glial cells: possible involvement of nitric oxide production from inducible nitric oxide synthase. *Glia* **22:**138-148; 1998.

[29] Mancuso, C.; Bonsignore, A.; Di Stasio, E.; Mordente, A.; Motterlini, R. Bilirubin and S-nitrosothiols interaction: evidence for a possible role of bilirubin as a scavenger of nitric oxide. *Biochem Pharmacol* **66:**2355-2363; 2003.

[30] Darwin, K. H.; Ehrt, S.; Gutierrez-Ramos, J. C.; Weich, N.; Nathan, C. F. The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. *Science* **302:**1963-1966; 2003.

[31] Venkatesh, J.; Kumar, P.; Krishna, P. S.; Manjunath, R.; Varshney, U. Importance of uracil DNA glycosylase in Pseudomonas aeruginosa and Mycobacterium smegmatis, G+C-rich bacteria, in mutation prevention, tolerance to acidified nitrite, and endurance in mouse macrophages. *J Biol Chem* **278:**24350-24358; 2003.

[32] Sambrook, J.; Russell, D. W. *Molecular Cloning-A Laboratory Manual.* Cold Spring Harbor, New York: Cold Spring Harbor laboratory Press; 2001.

[33] Cherepanov, P. P.; Wackernagel, W. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158:**9-14; 1995.

[34] Qiagen Inc. *The QIAexpresionist-A handbook for high-level expression and purification of 6xHis-tag proteins.* Valencia, CA: Qiagen Inc.; 2003.

[35] Pettersen, H. S.; Sundheim, O.; Gilljam, K. M.; Slupphaug, G.; Krokan, H. E.; Kavli, B. Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. *Nucleic Acids Res* **35:**3879-3892; 2007.

[36] Yao, M.; Hatahet, Z.; Melamede, R. J.; Kow, Y. W. Deoxyinosine 3' endonuclease, a novel deoxyinosine-specific endonuclease from Escherichia coli. *Ann N Y Acad Sci* **726:**315-316; 1994.

[37] Yao, M.; Kow, Y. W. Further characterization of Escherichia coli endonuclease V. *J Biol Chem* **272:**30774-30779; 1997.

[38] Weiss, B. Endonuclease V of Escherichia coli prevents mutations from nitrosative deamination during nitrate/nitrite respiration. *Mutat Res* **461:**301-309; 2001.

[39] Hitchcock, T. M.; Gao, H.; Cao, W. Cleavage of deoxyoxanosine-containing oligodeoxyribonucleotides by bacterial endonuclease V. *Nucleic Acids Res* **32:**4071-4080; 2004.

[40] Feng, H.; Dong, L.; Klutz, A. M.; Aghaebrahim, N.; Cao, W. Defining Amino Acid Residues Involved in DNA-Protein Interactions and Revelation of 3'-Exonuclease Activity in Endonuclease V. *Biochemistry* **44:**11486-11495; 2005.

[41] Buchmeier, N. A.; Lipps, C. J.; So, M. Y.; Heffron, F. Recombination-deficient mutants of Salmonella typhimurium are avirulent and sensitive to the oxidative burst of macrophages. *Mol Microbiol* **7:**933-936; 1993.

[42] Maragos, C. M.; Andrews, A. W.; Keefer, L. K.; Elespuru, R. K. Mutagenicity of glyceryl trinitrate (nitroglycerin) in Salmonella typhimurium. *Mutat Res* **298:**187-195; 1993.

[43] Tamir, S.; Burney, S.; Tannenbaum, S. R. DNA damage by nitric oxide. *Chem Res Toxicol* **9:**821-827; 1996.

[44] Darwin, K. H.; Nathan, C. F. Role for nucleotide excision repair in virulence of Mycobacterium tuberculosis. *Infect Immun* **73:**4581-4587; 2005.

[45] Laurent, T. C.; Moore, E. C.; Reichard, P. Enzymatic Synthesis of Deoxyribonucleotides. Iv. Isolation and Characterization of Thioredoxin, the Hydrogen Donor from Escherichia Coli B. *J Biol Chem* **239:**3436-3444; 1964.

[46] Holmgren, A. Thioredoxin. *Annu Rev Biochem* **54:**237-271; 1985.

[47] Powis, G.; Montfort, W. R. Properties and biological activities of thioredoxins. *Annu Rev Biophys Biomol Struct* **30:**421-455; 2001.

[48] Arner, E. S.; Holmgren, A. Physiological functions of thioredoxin and thioredoxin reductase. *Eur J Biochem* **267:**6102-6109; 2000.

[49] Miranda-Vizuete, A.; Damdimopoulos, A. E.; Gustafsson, J.; Spyrou, G. Cloning, expression, and characterization of a novel Escherichia coli thioredoxin. *J Biol Chem* **272:**30841-30847; 1997.

[50] Ritz, D.; Patel, H.; Doan, B.; Zheng, M.; Aslund, F.; Storz, G.; Beckwith, J. Thioredoxin 2 is involved in the oxidative stress response in Escherichia coli. *J Biol Chem* **275:**2505-2512; 2000.

[51] Prieto-Alamo, M. J.; Jurado, J.; Gallardo-Madueno, R.; Monje-Casas, F.; Holmgren, A.; Pueyo, C. Transcriptional regulation of glutaredoxin and thioredoxin pathways and related enzymes in response to oxidative stress. *J Biol Chem* **275:**13398-13405; 2000.

[52] Lillig, C. H.; Holmgren, A. Thioredoxin and related molecules--from biology to health and disease. *Antioxid Redox Signal* **9:**25-47; 2007.

[53] Ellgaard, L.; Ruddock, L. W. The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep* **6:**28-32; 2005.

[54] Knoblach, B.; Keller, B. O.; Groenendyk, J.; Aldred, S.; Zheng, J.; Lemire, B. D.; Li, L.; Michalak, M. ERp19 and ERp46, new members of the thioredoxin family of endoplasmic reticulum proteins. *Mol Cell Proteomics* **2:**1104-1119; 2003.

[55] Holmgren, A. Antioxidant function of thioredoxin and glutaredoxin systems. *Antioxid Redox Signal* **2:**811-820; 2000.

[56] Fernando, M. R.; Nanri, H.; Yoshitake, S.; Nagata-Kuno, K.; Minakami, S. Thioredoxin regenerates proteins inactivated by oxidative stress in endothelial cells. *Eur J Biochem* **209:**917-922; 1992.

[57] Spector, A.; Yan, G. Z.; Huang, R. R.; McDermott, M. J.; Gascoyne, P. R.; Pigiet, V. The effect of H2O2 upon thioredoxin-enriched lens epithelial cells. *J Biol Chem* **263:**4984-4990; 1988.

[58] Jeong, W.; Yoon, H. W.; Lee, S. R.; Rhee, S. G. Identification and characterization of TRP14, a thioredoxin-related protein of 14 kDa. New insights into the specificity of thioredoxin function. *J Biol Chem* **279:**3142-3150; 2004.

[59] Hirota, K.; Nakamura, H.; Masutani, H.; Yodoi, J. Thioredoxin superfamily and thioredoxin-inducing agents. *Ann N Y Acad Sci* **957:**189-199; 2002.

[60] Rhee, S. G.; Chae, H. Z.; Kim, K. Peroxiredoxins: a historical overview and speculative preview of novel mechanisms and emerging concepts in cell signaling. *Free Radic Biol Med* **38:**1543-1552; 2005.

[61] Wrammert, J.; Kallberg, E.; Leanderson, T. Identification of a novel thioredoxin-related protein, PC-TRP, which is preferentially expressed in plasma cells. *Eur J Immunol* **34:**137-146; 2004.

[62] Comtois, S. L.; Gidley, M. D.; Kelly, D. J. Role of the thioredoxin system and the thiol-peroxidases Tpx and Bcp in mediating resistance to oxidative and nitrosative stress in Helicobacter pylori. *Microbiology* **149:**121-129; 2003.

[63] Hess, D. T.; Matsumoto, A.; Kim, S. O.; Marshall, H. E.; Stamler, J. S. Protein S-nitrosylation: purview and parameters. *Nat Rev Mol Cell Biol* **6:**150-166; 2005.

[64] Benhar, M.; Forrester, M. T.; Stamler, J. S. Protein denitrosylation: enzymatic mechanisms and cellular functions. *Nat Rev Mol Cell Biol* **10:**721-732; 2009.

[65] Martinez, A.; Kolter, R. Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J Bacteriol* **179:**5188-5194; 1997.

Supplementary Content

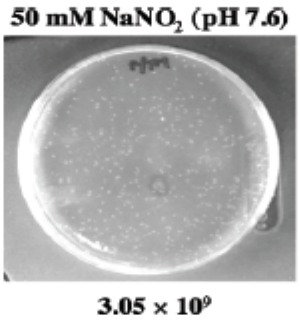

Figure S1. Effect of sodium nitrite on survival. E. coli BW1739 cells containing an empty pBluescript vector were treated with 50 mM sodium nitrite (pH 7.6) at 37°C for 10 min. Following incubation, a small portion (10 μl) of the cell mixtures were diluted into 1 ml of cold LB medium to stop the reaction and then diluted 2 additional times. The resulting dilution mixture (200 μl) was plated onto LB+AMP plate and incubated at 37°C for overnight.

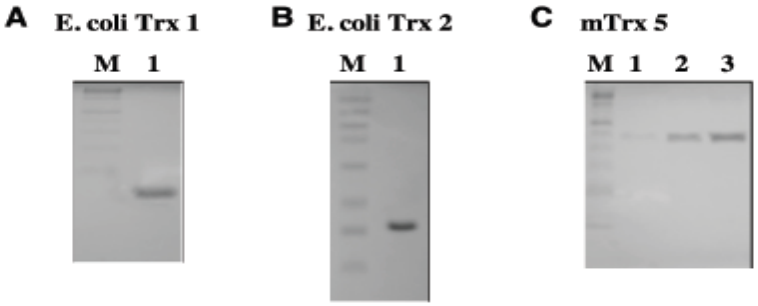

Figure S2. SDS-PAGE analysis. Purified proteins were electrophoresed on a 12.5 % polyacrylamide gel containing 0.1 % SDS. Protein bands were visualized by Coomassie blue staining. **A.** SDS-PAGE of E. coli thioredoxin 1. M, protein standard markers; lane 1: purified thioredoxin 1 protein. **B.** SDS-PAGE of E. coli thioredoxin 2. M, protein standard markers; lane 1: purified thioredoxin 2 protein. **C.** SDS-PAGE of mTrx5. M, protein standard markers; lanes 1-3: purified mTrx5 protein.

CHAPTER THREE

IDENTIFICATION OF *ESCHERICHIA COLI* MUG AS A ROBUST XANTHINE DNA GLYCOSYLASE

I. Abstract

The gene for the mismatch-specific uracil DNA glycosylase (MUG) was identified in the *Escherichia coli* genome as a sequence homolog of the human thymine DNA glycosylase (TDG) with activity against mismatched uracil base pairs. Examination of cell extracts led us to detect a previously unknown xanthine DNA glycosylase (XDG) activity in *E. coli*. DNA glycosylase assays with purified enzymes indicated the novel XDG activity is attributable to MUG. Here, we report a biochemical characterization of xanthine DNA glycosylase activity in MUG. The wild type MUG possesses more robust activity against xanthine than uracil and is active against all xanthine-containing DNA (C/X, T/X, G/X, A/X and single-stranded X). Analysis of potentials of mean force (PMF) indicates that the double-stranded xanthine base pairs have a relatively narrow energetic difference in base flipping while the tendency for uracil base flipping follows the order of C/U > G/U > T/U > A/U. Site-directed mutagenesis performed on conserved motifs revealed that N140 and S23 are important determinants for XDG activity in *E. coli* MUG. Molecular modeling and molecular dynamics simulations reveal distinct hydrogen bonding patterns in the active site of *E. coli* MUG, which account for the specificity differences between *E. coli* MUG and human thymine DNA glycosylase, as well as that between the wild type MUG and the N140 and S23 mutants. This study underscores the role of the favorable binding interactions in modulating the specificity of DNA glycosylases.

II. Introduction

DNA is constantly assaulted by environmental and endogenous agents, causing various types of chemical damage. DNA bases are subject to deamination by hydrolytic or oxidative reactions due to the reactivity of the exocyclic amino groups (1-3). Uracil (U), xanthine (X) and oxanine (O), hypoxanthine (I), and thymine (T) are the correspon ding deamination products derived from cytosine (C), guanine (G), adenine (A), and 5-methylcytosine, respectively. The amino-to-keto conversion of base deamination alters the hydrogen bond properties of the damaged bases from a hydrogen donor to a hydrogen bond acceptor, which may result in mutation during DNA replication.

Uracil DNA glycosylase (UDG), an enzyme present in organisms as simple as viruses or as complex as humans, initiates the repair of uracil in DNA. Five families, classified according to sequence and structural homologies, constitute a UDG super family (4,5). Family 1 includes the extensively studied Escherichia coli (E. coli), human and herpes simplex virus 1 UNGs. Family 2 contains human thymine DNA glycosylase (hTDG) and E. coli mismatch-specific uracil DNA glycosylase (MUG). hTDG is unique in its ability to excise thymine from a G/T mismatch generated from 5-methylcytosine deamination (6). Family 3 is comprised of SMUG1 (single-strand-selective monofun-ctional uracil DNA glycosylase) proteins found in vertebrates and some bacteria. Family 4 UDGs are iron-sulfur-containing enzymes found in prokaryotes. Family 5 are found in a limited number of species of prokaryotic organisms such as archaea. It is not uncom mon for an organism to possess more than one uracil DNA glycosylase. In addition to UDG, TDG and SMUG1, humans also have another uracil DNA glycosylase called

MBD4 that does not belong to the UDG superfamily (7).

Xanthine is now recognized as a stable lesion under physiological conditions (8,9). As such, although repair of xanthine was noted in human lymphoblast cells in an earlier study (10), enzymes that may repair xanthine were not identified until recently. Both E. coli AlkA and its functional homolog human alkyladenine DNA glycosylase (hAAG) have xanthine DNA glycosylase activity (9,11,12). The SMUG1 enzymes from bacteria and humans are also active on xanthine-containing DNA (13). In addition, several homologs of bacterial endonuclease V exhibit deoxyxanthosine endonuclease activity (11,14-16).

Identification of DNA repair activity in E. coli has led to the discovery of new repair enzymes or novel activities. DNA repair deficient mutant strains have facilitated identification of functional homologs in eukaryotic systems (17-20). We are interested in achieving a comprehensive understanding of xanthine DNA repair in E. coli. Previous studies show that AlkA, endo VIII, and endo V in E. coli possess xanthine repair acti vities (11,12,16). Using an E. coli triple mutant strain (nfi nei alkA), we detected xanthine DNA glycosylase (XDG) activity in whole cell protein extracts. Further bioch emical analysis led to the discovery of XDG activity in the mismatch-specific uracil DNA glycosylase MUG. Surprisingly, kinetic analysis revealed that the XDG activity from MUG was more robust than UDG activity. Rather than being active with only double-stranded mismatch uracil base pairs, MUG can excise xanthine from double-stranded base pairs as well as single-stranded DNA. Structural elements that are involved in determining the base recognition in MUG were probed by site-directed mutagenesis.

Mutational effects on glycosylase activities of deaminated bases, X and U, were analyzed by activity assays and binding analyses. Thermodynamic properties associated with the flipping of deaminated base pairs were determined by calculating potentials of mean force. The base recognition specificity is discussed in light of molecular modeling and molecular dynamics simulations of MUG interactions with deaminated bases.

III. Experimental Procedures

A. Preparation of *E. coli* Cell Extracts

Bacterial cells (BW1466 and BW1739) from 100-ml cultures grown to late exponential phase were harvested by centrifugation at 5,000 rpm with GSA-10 rotor in RC5C Sorvall centrifuge (Dupont). The cell pellets were suspended in 5 ml sonication buffer and sonicated 5 times with a burst duration of 1 min each. The lysates were centrifuged at 12,000 rpm at 4 °C for 20 min. The supernatants containing soluble proteins were transferred to fresh tubes, filtered with 0.45 μm syringe filters (Whatman, Clifton, New Jersey) and dialyzed at 4 °C overnight against a buffer containing 20 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), and 0.1 mM DTT. Protein concentrations were measured by the Bradford method using bovine serum albumin as a standard (21).

B. Oligodeoxynucleotide Substrates

The fluorescently labeled oligodeoxynucleotide substrates were prepared as described (22). The sequences of the oligonucleotides are shown in Fig. 3.1A. Oligo deoxyribonucleotides were ordered from IDT, purified by PAGE, and dissolved in TE buffer at a final concentration of 10 μM. The two complementary strands with the unlabeled strand in 1.2-fold molar excess were mixed, incubated at 85 °C for 3 min, and

allowed to form duplex DNA substrates at room temperature for more than 30 min. The xanthine- and oxanine-containing oliogonucleotide were constructed as previously des cribed (15,23).

C. DNA Glycosylase Activity Assay

DNA glycosylase cleavage assays for *E. coli* MUG were performed at 37 °C for 60 min in a 10 µl reaction mixture containing 10 nM oligonucleotide substrate, an indicated amount of glycosylase protein, 20 mM Tris-HCl (pH 7.5), 100 mM KCl, 5 mM EDTA, and 2 mM 2-mercaptoethanol. The resulting abasic sites were cleaved by in cubation at 95 °C for 5 min after adding 0.5 µl of 1 N NaOH. Reactions were quenched by addition of an equal volume of GeneScan stop buffer. After incubation at 95 °C for 3 min, samples (3.5 µl) were loaded onto a 7 M urea-10% denaturing polyacrylamide gel. Electrophoresis was conducted at 1500 V for 1.5 h using an ABI 377 sequencer (Applied Biosystems). Cleavage products and remaining substrates were quantified using Gene Scan analysis software.

D. Gel Mobility Shift Assay

The binding reactions were performed on ice for 10 min in a 10-µl volume containing 50 nM DNA substrate, 20 mM Tris-HCl (pH 7.2), 50 mM NaCl, 5 mM EDTA, 1 mM DTT, 0.1 mg/ml bovine serum albumin, 10% glycerol, and the indicated amount of *E. coli* MUG protein. Samples were supplemented with 2 µl of 100% glycerol and electrophoresed at 200 V on a 6% native polyacrylamide gel in 1 x TB buffer (89 mM Tris base and 89 mM boric acid) supplemented with 5 mM EDTA. The bound and free DNA species were analyzed using a Typhoon 9400 Imager (Molecular Dynamics)

with the following settings: photomultiplier tube at 600 V, excitation at 495 nm, and emission at 535 nm.

E. Molecular Modeling and Molecular Dynamics Simulations

Molecular models of the unbound and bound conformations of wild type (wt) *E. coli* MUG were used as initial structures for subsequent computational analyses. The crystal structure of *E. coli* MUG (pdb accession code 1mug) was used as a model for the unbound MUG enzyme. The molecular model of the wt *E. coli* MUG complexed with a DNA decamer sequence containing uracil was constructed based on the crystal structure of UDG bound to a DNA decamer (pdb accession code 1emh).

Molecular dynamics simulations were performed on the bound MUG structures using the CHARMM 32b1 molecular mechanics software package (24), and the CHARMM 27 force field (25,26). Interaction energies consisting of Coulomb and van der Waals potential energies were calculated over the molecular dynamics trajectory between the active site residues and the substrates using the "coor inter" module in CHARMM.

Potentials of mean force (PMF) describe free energy changes along a pre-defined reaction coordinate while averaging over the remaining degrees of freedom. Here, potentials of mean force are used to describe the free energy changes associated with rotating a nucleotide from the interior of the DNA double helix into the aqueous solvent. A detailed description of the computational methods is provided in the Supplemental Data.

IV. Results

A. Detection of XDG Activity in *E. coli* Cell Extracts

Studies of uracil repair in *E. coli* has led to the discovery of a uracil DNA glycosylase and a mismatch-specific uracil DNA glycosylase (MUG). Previous

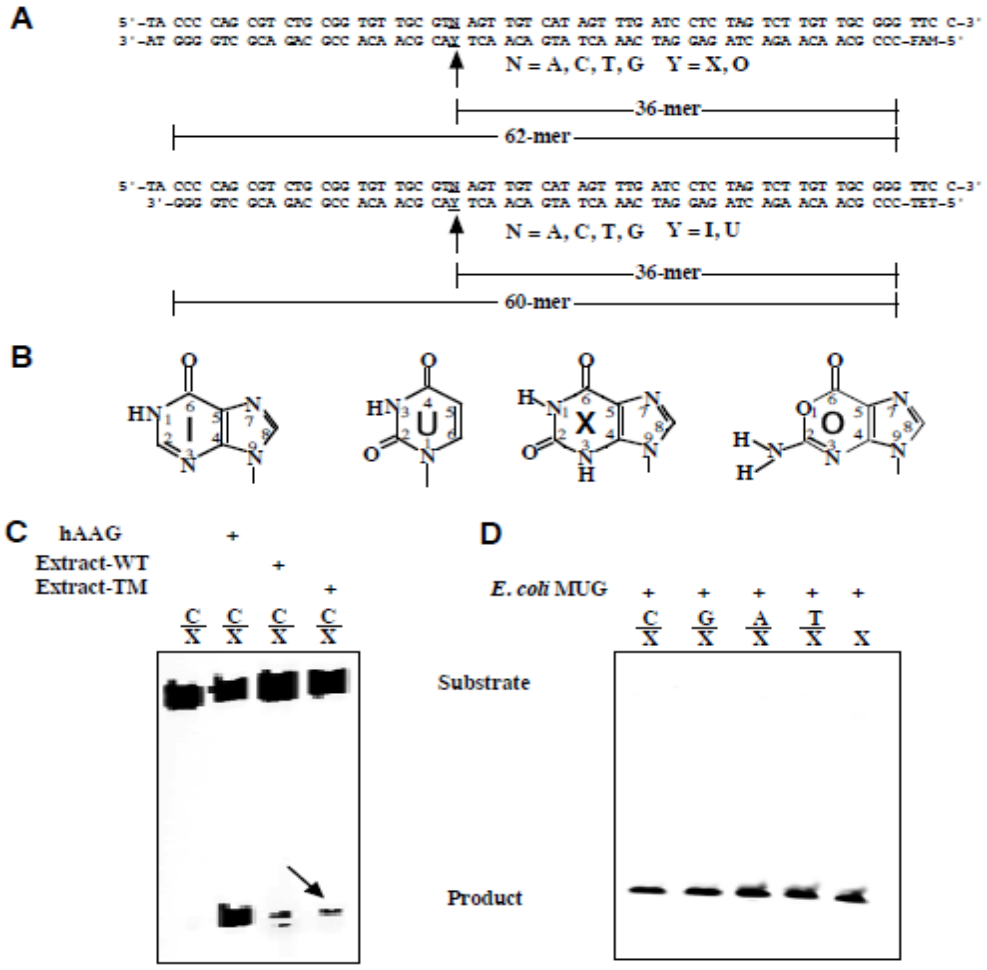

Figure 3. 1 Cleavage of deaminated base-containing DNA substrates by *E. coli* cell extracts and *E. coli* MUG from a commercial source. Cleavege reactions were performed as described in Experimental Procedures with 5 µl cell extract or 1 unit of wt MUG protein (Trevigen) and 10 nM substrate. **A.** Sequences of xanthine (X)- and oxanine (O)-, hypoxanthine (I)- and uracil (U)-containing oligodeoxyribonucleotide substrates. **B.** Chemical structures of deaminated DNA bases. **C.** DNA glycosylase activity on C/X in *E. coli* cell extracts. **D.** DNA glycosylase activity of MUG on X-containing substrate. WT: BW1466; TM: triple mutant, BW1739. hAAG was assayed as a control with 100 nM hAAG protein and 10 nM substrate as described previously (23). Under the assay conditions, the substrate was completely cleaved to product.

investigations using purified enzymes revealed that two glycosylases (AlkA and endo VIII) and endo V in *E. coli* possess xanthine DNA glycosylase activity or deoxy-xanthosine endonuclease activity (11,12,16). However, XDG activities from the two glycosylases, particularly the latter, are quite low. To survey whether *E. coli* contains additional XDG enzymes, we examined the cleavage activity in cell extracts of a triple mutant strain which had *alkA*, *nei* (endo VIII), and *nfi* (endo V) deleted. The assays were performed using fluorescently labeled xanthine-containing deoxyoligonucleotide substrates (Fig. 3.1A-B). As expected, we detected cleavage of C/X in the wild type cell extract (Fig. 3.1C). Surprisingly, we also detected xanthine DNA glycosylase activity in the triple mutant strain that eliminated the activity of all previously known XDG enzymes (Fig. 3.1C). This result indicated that the *E. coli* genome contained an additional XDG enzyme. Given that a previous study had investigated XDG activity in purified *E. coli* AlkA, endo III, endo V, endo VIII, Fpg/MutM, MutY and UDG (11), we surmised that MUG, which was not included in the previous study, may be accountable for the observed XDG activity in the triple mutant cell extract. A quick test was performed using purified MUG from a commercial source (Trevigen). As shown in Fig. 1D, MUG was found to be active on all five xanthine-containing substrates, including the single-stranded substrate. We also tested XDG activity using a MUG single deletion strain and a quadruple deletion strain. XDG activity was not detected in either case, indicating that MUG is the predominant activity in cell extracts.

B. XDG Activity in Purified *E. coli* MUG

To further confirm the novel XDG activity, we set out to clone, express, and purify the wt MUG and an active site mutant MUG-N18A. MUG was expressed in the *ung mug* knockout strain BH214 to avoid contaminating activity from the host.

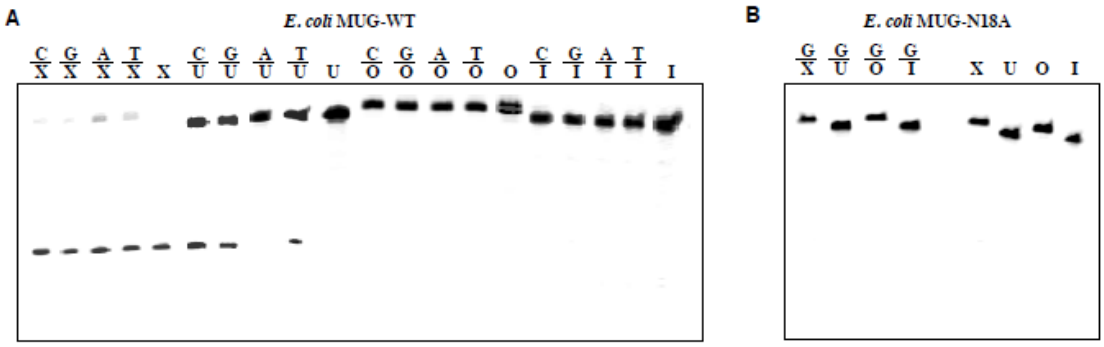

Figure 3.2 Cleavage of xanthine (X)-, uracil (U)-, oxanine (O)- and hypoxanthine (I)- containing DNA substrates by *E. coli* MUG. Cleavege reactions were performed as described in Experimental Procedures with 100 nM wt *E. coli* MUG protein and 10 nM substrate. **A.** Cleavage by wt *E. coli* MUG. **B.** Cleavage by N18A mutant of *E. coli* MUG.

We tested deaminated base repair activity on X-, U-, O-, and I-containing substrates (Fig. 3.2A). The purified wt MUG again showed cleavage of all five X-containing substrates (Fig. 3.2A). The uracil DNA glycosylase activity observed in the wt MUG was essentially identical to that reported in previous studies, i.e., active on C/U, G/U and T/U mismatched uracil base pairs (27,28). The XDG activity was noticeably more robust than the UDG activity as indicated by close-to-complete cleavage of the xanthine substrates (Fig. 3.2A). No oxanine or hypoxanthine DNA glycosylase activities were detected from the wt enzyme under our assay conditions (Fig. 3.2A). To further verify that the xanthine DNA glycosylase activity was authentic to MUG, we performed the same assay using an active site mutant MUG-N18A, which abolished the ability of the enzyme to activate a water molecule to attack the N-glycosidic bond (29,30). No glycosylase activity on

deaminated bases was detected, confirming that the XDG activity was native to the MUG protein (Fig. 3.2B).

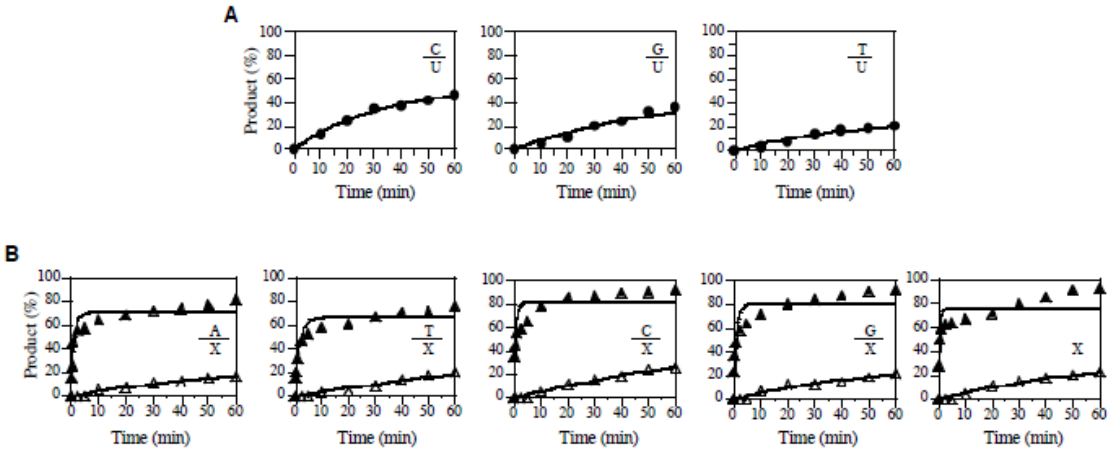

Figure 3.3 Kinetic analysis of glycosylase activity of wt *E. coli* MUG on X- and U-containing substrates. A. Time course analysis of cleavage activity on U-containing substrates. Cleavege reactions were performed as described in Experimental Procedures with 100 nM wt *E. coli* MUG protein and 10 nM substrate. E:S = 10:1. **B.** Time course analysis of cleavage activity on X-containing substrates. S = 10 nM. (⑩) E:S = 10:1; (△) E:S = 1:10.

The robust XDG activity in MUG prompted us to quantitatively determine the deaminated base repair activity in MUG. Under the condition that the enzyme was in excess (E:S ratio = 10:1), removal of uracil in the three mismatched base pairs was less than 50% (Fig. 3.3A). Alternatively, excision of xanthine is significantly more efficient, resulting in a close-to-complete cleavage (Fig. 3.3B). The apparent rate constants were in general comparable among the double-stranded xanthine substrates but appeared to be slightly more active on C/X and G/X substrates (Table 3.1). However, MUG exhibited a much more robust DNA glycosylase activity on xanthine-containing DNA than uracil-containing DNA. For example, the cleavage efficiency of C/X was about 13-fold higher than that of C/U (Table 3.1). Interestingly, the XDG activity on C/X base pair, the likely

biological substrate resulting from the direct deamination of a guanine base, is slightly stronger than other base pairs. Within the uracil-containing substrates, MUG is most

Table 3.1 Apparent rate constants for cleavage of xanthine (X) and uracil (U) substrates by *E. coli* MUG and mutants (min^{-1})[a]

MUG	Bottom Strand	Top Strand				Single Strand
		C	G	A	T	
WT	X	0.24	0.23	0.21	0.21	0.36
I17L		0.024	0.016	0.013	0.015	0.022
S23A		0.071	0.046	0.024	0.027	0.059
L144S		0.012	0.010	0.011	0.011	0.012
WT	U	0.018	0.014	n.a.[b]	0.0094	n.a.
S23A		0.032	0.031	n.a.	0.024	n.a.
N140H		0.0029	0.0021	n.a.	0.00034	n.a.

[a]: The reactions were performed as described in Experimental Procedures with 100 nM MUG and 10 nM substrates. Data are an average of two independent experiments.
[b]: n.a.: no activity was detected under assay conditions.

active on C/U followed by G/U and then by T/U (Table 3.1). No activity was detected on A/U or single-stranded U (Fig. 3.2 and Table 3.1). Under the condition that the substrate was in excess (E:S ratio = 1:10), cleavage of xanthine-containing substrates reached a level of approximately 20% (Fig. 3.3B). However, no cleavage of uracil-containing substrates was detected (data not shown). Although MUG was discovered as a uracil DNA glycosylase, these results suggested that MUG was more efficient as an XDG than as a UDG.

C. Identification of XDG Activity Determinants

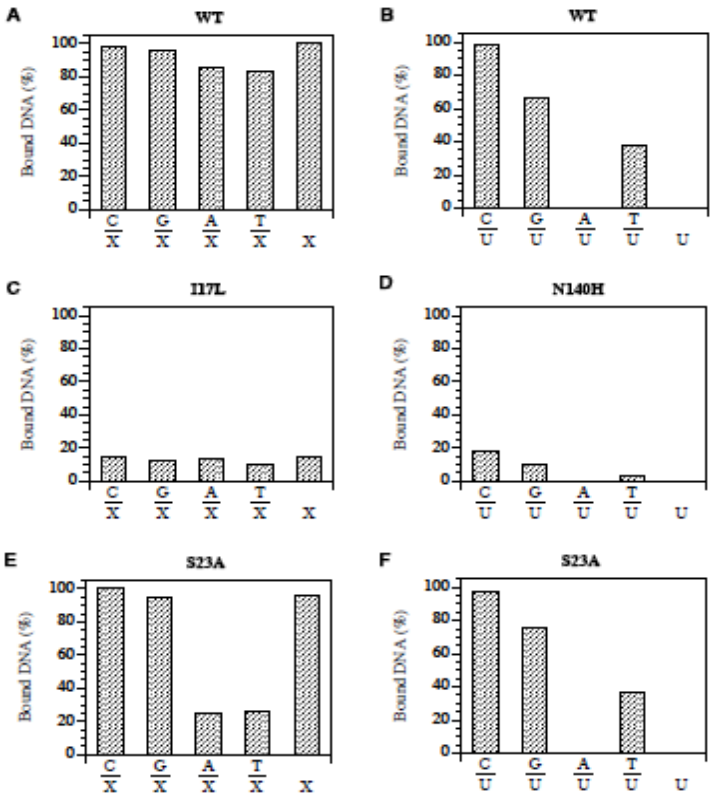

Figure 3.4 Binding of X-, U-containing DNA substrates by *E. coli* MUG. Gel mobility shift analysis was performed as described in Experimental Procedures with 500 nM protein and 50 nM substrate. Data are an average of two independent experiments. **A.** wt MUG with X-containing DNA. **B.** wt MUG with U-containing DNA. **C.** I17L with X-containing DNA. **D.** N140H with U-containing DNA. **E.** S23A with X-containing DNA. **F.** S23A with U-containing DNA.

To identify amino acid residues that may play a role in recognition of deaminated bases, we selected eight positions, in motifs I and II that define the base recognition pocket, for a site-directed mutagenesis study (Fig. S1). Six positions are located in motif 1 and two positions are located in motif 2. S23 and N140 were identified as major determinants of the XDG activities.

Alanine substitution at the S163 position in *Schizosaccharomyces pombe* TDG results in the complete loss of XDG activity (31). The same substitution in the equivalent S23 position in *E. coli* MUG exhibited an interesting effect. The UDG activity on C/U,

G/U and T/U substrates was enhanced (Fig. S2-B and Table 3.1). On the other hand, the XDG activity was reduced, ranging from 3-fold for C/X, to 5-fold for G/X, to 9-fold for A/X, to 8-fold for T/X, and to 6-fold for single-stranded X-containing DNA (Fig. S2-B and Table 3.1). The role that S23 may play in xanthine recognition is discussed later in light of modeled MUG-X complex structure.

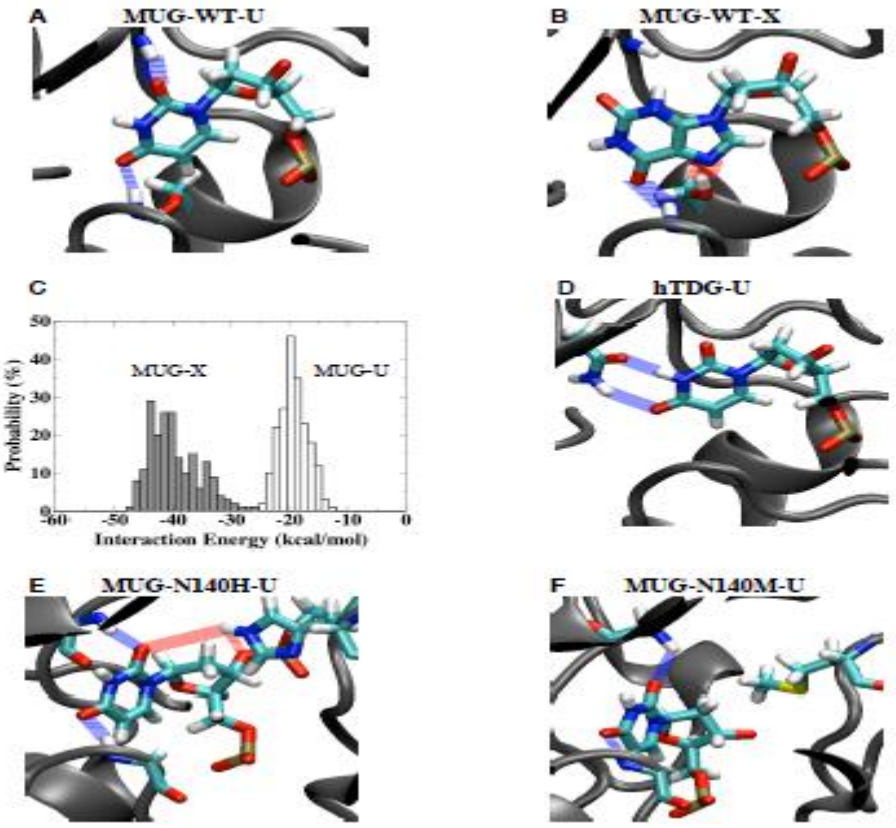

Figure 3.5 Molecular modeling of base recogition by *E. coli* MUG. A. Interactions between wt *E. coli* MUG and uracil. Mainchain hydrogen bonding between N18, F30 and uracil are shown in blue. **B.** Interactions between wt *E. coli* MUG and xanthine. Mainchain hydrogen bonding between F30 and uracil is shown in blue. Sidechain hydrogen bonding between S23 and N^7 of xanthine is shown in red. **C.** Energetics of wt *E. coli* MUG interactions with G/X (solid bars) and G/U base pairs (blank bars). **D.** Interactions between human TDG and uracil. Sidechain hydrogen bonding between N191 and uracil are shown in blue. **E.** Interactions between *E. coli* MUG-N140H and uracil. Hydrogen bonding between the sidechain of N140H and the uracil and that of the 3'-phosphate are shown in red. Mainchain hydrogen bonding between N18, F30 and uracil are shown in blue. **F.** Interactions between *E. coli* MUG-N140M and uracil. Mainchain hydrogen bonding between N18, F30 and uracil are shown in blue.

Two mutations were constructed at the N140 position of motif 2. N140M, which mimicked human TDG, resulted in a complete loss of XDG and UDG activity (data not shown). N140H, which mimicked family 1 UDGs, only maintained a reduced UDG activity toward C/U, G/U, and a substantially reduced activity towards T/U (Fig. S2-C and Table 3.1). Strikingly, no xanthine DNA glycosylase activity was detected in N140H under the assay conditions (Fig. S2-C). The N140 position in motif 2 is a His residue in all UDG superfamily proteins except for the MUG/TDG family (Fig. S1). This His residue forms a hydrogen bond with the C^2-carbonyl oxygen of uracil in UNG and SMUG1 (32-34). However, in the MUG structure, N140 is too far (>4Å) to make a direct contact with the C^2-carbonyl oxygen (30). The effect of substitutions at the N140 position was further studied by molecular modeling as described later.

In addition, I17L and L144S had a similar effect, i.e., a significant reduction in XDG activity and a loss of UDG activity (Fig. S2 and Table 3.1). A detailed description of these two and other mutants is provided in the Supplemental Data.

D. Binding Analysis

To better understand the mutational effects on the binding affinities to deaminated base-containing substrates, we performed gel mobility shift analyses. Consistent with the results obtained through the activity assays, the wt MUG interacted with all xanthine-containing substrates and C/U, G/U and T/U to form a stable complex (Fig. 3.4A-B). In keeping with its glycosylase activity, I17L was able to bind to all xanthine-containing substrates but not any uracil-containing substrates (Fig. 3.4C and data not shown). For S23A, the binding to uracil-containing substrates remained similar to the wt enzyme;

however, binding to some xanthine-containing substrates such as A/X and T/X was reduced by approximately 4-fold (Fig. 3.4E-F). Consistent with the loss of XDG activity

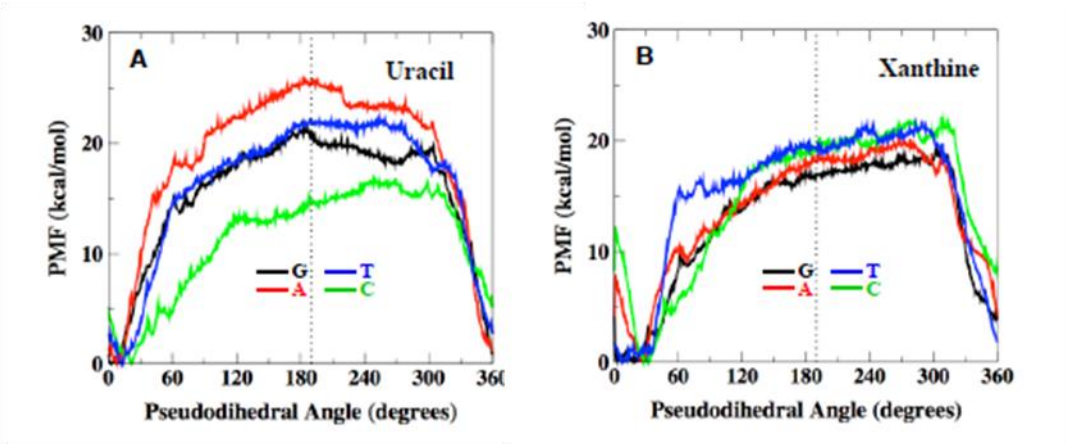

Figure 3.6 Potentials of mean force (PMF) of uracil- and xanthine-containing base pairs along the pseudodihedral angle coordinate. Watson-Crick base pairing is approximately 10°-30° pseudodihedral angle and the flipped out state is approximately 190°. **A.** Uracil-containing base pairs. **B.** Xanthine-containing base pairs.

and the reduction of UDG activity, N140H showed no detectable binding affinity towards all xanthine-containing substrates and much reduced affinity towards substrates containing C/U, G/U and T/U base pairs (Fig. 3.4D and data not shown). N140M showed no binding to either X- or U-containing substrates (data not shown).

V. Discussion

A. MUG as a Xanthine DNA Glycosylase

E. coli MUG was initially discovered as a mismatch-specific double-stranded uracil DNA glycosylase that shared significant sequence homology with the glycosylase domain of human TDG (35). Its substrate specificity has been broadened to include 3,N^4-ethenocytosine and 8-(hydroxymethyl)-3,N^4-ethenocytosine (28,36), thymine (28,29), 5-hydroxymethyluracil (37) and more interestingly the guanine derivative 1,N^2-

ethenoguanine (38). Xanthine, as a deaminated product of guanine, can be viewed structurally as a fusion of a uracil ring with an imidazole ring (Fig. 3.1B). Our biochemical analysis using *E. coli* cell extracts and purified proteins provides evidence that MUG can also act as a xanthine DNA glycosylase. Notably different from other glycosylase activities, MUG excises xanthine from both single-stranded and double-stranded DNA. It was reported that MUG was more effective in removing 3,N^4-ethenocytosine than uracil (28). However, MUG did not seem to be active on single-stranded 3,N^4-ethenocytosine-containing DNA (28). Another surprising finding is that the efficiency of removing xanthine is at least 10-fold greater than that of removing uracil (Table 3.1), suggesting that MUG is more robust as an XDG than a UDG. These results reinforce the question of whether the biological role MUG plays in a cell is more than just a uracil repair enzyme, rather, it may help repair xanthine or other damaged bases as well (5,27,28,39). While human TDG does not have XDG activity, in contrast to *E. coli* MUG, mammalian systems may utilize alkyladenine DNA glycosylase (AAG), SMUG1 or other DNA glycosylases to excise xanthine (9,13).

Given the highly robust xanthine DNA glycosylase activity, we set out to understand how the active site of *E. coli* MUG accommodates a xanthine base. Molecular models were constructed to characterize how a xanthine base may fit into the active site of *E. coli* MUG. Both uracil and xanthine are accommodated in the active site without significant distortion of the enzyme structure (Fig. 3.5A-B). Uracil forms two hydrogen bonds with the mainchains of N18 and F30 (Fig. 3.5A). On the other hand, xanthine is stabilized by a mainchain interaction with F30 and a sidechain interaction

with S23 (Fig. 3.5B). The sidechain interaction of S23 with N^7 of xanthine apparently plays a role in xanthine recognition. Alanine substitution at the S23 position, which eliminates the sidechain interaction, resulted in reduced binding affinity to A/X and T/X substrates (Fig. 3.4E). More profoundly, the XDG activity was also reduced, in particular with A/X and T/X substrates (Fig. S2-B and Table 3.1). Although substitutions at the Ser residue in MUG does not cause a complete loss of XDG activity as seen in *S. pombe* TDG (31), the sidechain of S23 in MUG appears to provide a favorable interaction that facilitates the recognition of xanthine base.

To further understand the interactions of MUG in the bound state, structural ensembles were constructed for MUG bound to uracil and xanthine through molecular dynamics simulations. Electrostatic and van der Waals interaction energies between the deaminated base (xanthine or uracil) and active site residues (within ~8Å of substrate) were calculated over the 400 ps of production trajectory. Based on these calculations, MUG is capable of interacting favorably in the bound state with the xanthine substrate (Fig. 3.5C). The average interaction energies calculated between MUG and xanthine are stronger than those between MUG and uracil (Fig. 3.5C). The ability to accommodate xanthine is consistent with previous studies, which show that ethenocytosine derivatives can comfortably fit into the binding pocket (29,36). However, the activity on deaminated purine bases seems to be limited to the xanthine base. The favorable interactions with xanthine as shown in Fig. 3.5C may determine its specificity as a xanthine DNA glycosylase.

B. Comparison of *E. coli* MUG and Human TDG

To understand the structural differences that may underlie the functional distinction, we created bound models of hTDG to compare the differences between how hTDG interacts with uracil and xanthine. While uracil was stabilized by sidechain interactions provided by N191 (Fig. 3.5D), xanthine appeared to have fewer favorable interactions in the active site (data not shown). The favorable sidechain hydrogen bonding between S23 of MUG and xanthine is not available because the position is occupied by an Ala residue (A145) in hTDG (Fig. S1). The reduction of XDG activity observed in MUG-S23A mutant illustrates the role of this interaction in xanthine recognition.

To further investigate the different activities in the wt MUG and S23A mutant, differences in protein-DNA interaction energies were examined within canonical ensem bles generated using molecular dynamics. Unsurprisingly, the results indicate that the wt MUG has stronger electrostatic and van der Waals interactions with xanthine than the S23A mutant, as a result of the loss of a hydrogen bond between the sidechain hydroxyl of S23 and N^7 in xanthine (Fig. S3-A). Interestingly, the S23A mutant has a higher level of catalytic activity against C/U, G/U and T/U base pairs than the wt enzyme (Table 3.1). This is also consistent with the outcome of MD simulation, indicating a stronger inter action between S23A and uracil (Fig. S3-B). To better understand the origin of the enhancement, we performed per residue decomposition of the interaction energies. It appears that the mainchains of F30 and N18 form stronger hydrogen bonds with uracil in the S23A mutant (Fig. S3-C). To examine the impact of the Ala substitution at the S23 position on the dynamic motion of the protein, we calculated the mean square fluctuation

differences (ΔMSF) between the wt MUG and the S23A mutant. One change observed during the analysis was that motif 2 became more rigid in the S23A mutant (Fig. S3-D). Given that motif 2 provides a wedge to occupy the space vacated by the flipped base, it is possible that a more rigid wedge can be more effective in keeping the damaged base in a flipped out conformation. This could be the result of a reduction in the entropic penalty, resulting from the wedge interaction with the DNA. These analyses are consistent with the experiments showing that the S23A mutant has a higher level of catalytic activity against C/U, G/U and T/U base pairs than the wt enzyme (Table 3.1).

C. N140 and Xanthine DNA Glycosylase Activity

The role of individual amino acids in base recognition was probed by site-directed mutagenesis. Most of the mutants still maintain activity on xanthine-containing DNA. In a stark contrast, two of the N140 mutants we constructed (N140M and N140H) showed no detectable XDG activity (Fig. S2-C and data not shown). Substitution with Met also results in the loss of UDG activity while substitution with His reduces the UDG activity. Given that the wt MUG is much more robust on xanthine than uracil, the complete loss of XDG activity while still maintaining some UDG activity is dramatic (Fig. 3.2 and Fig. S2-C). These data underscore the role that N140 may play in modulating XDG activity. Molecular models of the N140 mutants bound to uracil and xanthine were constructed in order to investigate the interactions at position 140. In the modeled MUG-uracil complex structure, N140 in MUG interacts with the phosphate backbone through hydrogen bonding (Fig. S4-A), which may enhance the DNA binding. Although N140 in MUG is sequentially aligned with M269 in hTDG, the structural alignment of these enzymes,

performed with SPDBV (40), superimposes N140 of MUG with S271 of hTDG. Likewise, S271 of hTDG could form equivalent hydrogen bonds with the phosphate backbone (Fig. S4-B). In the modeled N140H-uracil structure, N140H appears capable of forming a hydrogen bond with the C^2-keto of uracil and a weak hydrogen bond with the 3'-phosphate (Fig. 3.5E). The presence of these favorable interactions may underscore the weak but observable UDG activity of the N140H mutant (Fig. S2-C and Fig. 3.4D). However, these potential interactions are lost when the uracil is substituted by xanthine (Fig. S4-C), which may explain the loss of XDG activity. The loss of both XDG and UDG activity in N140M can be viewed as due to the loss of DNA backbone interactions as seen in N140 of MUG and S271 of hTDG or loss of direct hydrogen bonding to uracil as seen in N140H. The lack of favorable interactions with the backbone or the base may lead to the complete loss of both XDG and UDG activity (Fig. 3.5F).

D. Base Pair Stability and DNA Repair Activity

An obvious difference between the XDG and UDG activity on the double-stranded DNA is that *E. coli* MUG is active on all xanthine-containing DNA but only active on C/U, G/U, and T/U substrates (Figs. 3.2-3.3). It is known that mismatched uracil-containing base pairs such as G/U are thermodynamically less stable than the A/U base pair (41). However, data on the stability of xanthine-containing base pairs are limited (42). To understand whether the difference in activity is related to conformational stability of the DNA, the thermodynamic stabilities of xanthine- and uracil-containing base pairs were determined by constructing the corresponding potential

of mean force (PMF) profiles of the base flipping mechanism. These PMF's provide a clear description of the thermodynamic tendencies associated with DNA base flipping.

The potentials of mean force generated through umbrella sampling indicate a greater thermodynamic tendency for deaminated bases to flip out of an isolated B-form DNA double helix relative to undamaged bases. Considering uracil flipping first, the calculated potentials of mean force demonstrate that the thermodynamic tendency for uracil to flip is significantly reduced when paired with adenine (Fig. 3.6A). This has been noted previously and is expected given that uracil forms a stable Watson-Crick base pair with adenine through two hydrogen bonds (43). Results further indicate that uracil has the greatest tendency to flip when paired with cytosine, followed by G/U and T/U base pairs (Fig. 3.6A). Interestingly, the PMF data are quite consistent with the UDG activity profile reported here (Fig. 3.3A). These results indicate that the tendency of the mismatched uracil-containing base pairs to flip out of the helix greatly facilitates their recognition by *E. coli* MUG. In contrast, the umbrella sampling results indicate that the paired base has relatively little influence on the thermodynamic tendency of xanthine to flip out of the DNA. As a consequence, the four xanthine-containing base pairs show a relatively narrow difference in the free energy of flipping (Fig. 3.6B). The XDG activity profile is, in general, consistent with similar flipping tendencies of the xanthine-containing base pairs (Fig. 3.3B). The role that poor base stacking and base flipping may play in DNA lesion recognition has been discussed recently (44-46). The data presented here on the wt *E. coli* MUG are in general in accord with the hypothesis that spontaneous base flipping plays a role in determining the catalytic efficiency. However, one should

keep in mind that how a glycosylase interacts with the damaged base in the base recognition pocket or the wedge region will also influence the catalytic efficiency of the base removal.

In summary, this work reports that MUG is a robust xanthine DNA glycosylase despite the fact that it is generally considered as a uracil DNA glycosylase in the extensively studied organism *E. coli*. The correlation of the activity profiles with base flipping energetics underlies the role of spontaneous base flipping in initial damaged base recognition and subsequent catalysis. The ability to recognize both a deaminated pyrimidine base and a purine base underscores the plasticity of the active site, a feature that distinguishes *E. coli* MUG from human TDG in the same family and UNGs in family 1 of the UDG superfamily. The ability to favorably interact with a DNA base lesion provides a means to determine the specificity of DNA glycosylases.

VI. Footnotes

This project was supported in part by Cooperative State Research, Education, and Extension Service/United States Department of Agriculture (SC-1700274, technical contribution No. 5805), Department of Defense (W911NF-05-1-0335, W911NF-07-1-0141, W81XWH-10-1-0385 to W. Cao), the National Institutes of Health (GM090141 to W. Cao), the National Science Foundation (MCB-0953783 to B. N. Dominy) and Howard Hughes Medical Institute-SC-Life. We thank Dr. Bernard Weiss at Emory University, Dr. Ashok Bhagwat at Wayne State University and Dr. Richard Cunningham at State University of New York for providing *E. coli* strains.

VII. References

1. Lindahl, T. (1993) *Nature* **362**, 709-715.
2. Shapiro, R. (1981) Damage to DNA caused by hydrolysis. in *Chromosome Damage and Repair* (Seeberg, E., and Kleppe, K. eds.), Plenum Press, New York
3. Suzuki, T., Yamaoka, R., Nishi, M., Ide, H., and Makino, K. (1996) *J. Am. Chem. Soc.* **118**, 2515-2516
4. Huffman, J. L., Sundheim, O., and Tainer, J. A. (2005) *Mutat Res* **577**, 55-76
5. Pearl, L. H. (2000) *Mutat Res* **460**, 165-181
6. Cortazar, D., Kunz, C., Saito, Y., Steinacher, R., and Schar, P. (2007) *DNA Repair (Amst)* **6**, 489-504
7. Hendrich, B., Hardeland, U., Ng, H. H., Jiricny, J., and Bird, A. (1999) *Nature* **401**, 301-304
8. Vongchampa, V., Dong, M., Gingipalli, L., and Dedon, P. (2003) *Nucleic Acids Res* **31**, 1045-1051
9. Wuenschell, G. E., O'Connor, T. R., and Termini, J. (2003) *Biochemistry* **42**, 3608-3616
10. Jaruga, P., and Dizdaroglu, M. (1996) *Nucleic Acids Res* **24**, 1389-1394
11. Dong, M., Vongchampa, V., Gingipalli, L., Cloutier, J. F., Kow, Y. W., O'Connor, T., and Dedon, P. C. (2006) *Mutat Res* **594**, 120-134
12. Terato, H., Masaoka, A., Asagoshi, K., Honsho, A., Ohyama, Y., Suzuki, T., Yamada, M., Makino, K., Yamamoto, K., and Ide, H. (2002) *Nucleic Acids Res* **30**, 4975-4984

13. Mi, R., Dong, L., Kaulgud, T., Hackett, K. W., Dominy, B. N., and Cao, W. (2009) *J Mol Biol* **385**, 761-778

14. Feng, H., Dong, L., Klutz, A. M., Aghaebrahim, N., and Cao, W. (2005) *Biochemistry* **44**, 11486-11495

15. Feng, H., Klutz, A. M., and Cao, W. (2005) *Biochemistry* **44**, 675-683

16. He, B., Qing, H., and Kow, Y. W. (2000) *Mutat Res* **459**, 109-114

17. Chakravarti, D., Ibeanu, G. C., Tano, K., and Mitra, S. (1991) *J Biol Chem* **266**, 15710-15715

18. Chen, J., Derfler, B., Maskati, A., and Samson, L. (1989) *Proc Natl Acad Sci U S A* **86**, 7961-7965

19. Samson, L., Derfler, B., Boosalis, M., and Call, K. (1991) *Proc Natl Acad Sci U S A* **88**, 9127-9131

20. van der Kemp, P. A., Thomas, D., Barbey, R., de Oliveira, R., and Boiteux, S. (1996) *Proc Natl Acad Sci U S A* **93**, 5197-5202

21. Coligan, J. E., Dunn, B. M., Ploegh, H. L., Speicher, D. W., and Wingfield, P. T. (1998) *Current Protocols in Protein Science*, John Wiley & Sons, Inc., New York

22. Huang, J., Lu, J., Barany, F., and Cao, W. (2001) *Biochemistry* **40**, 8738-8748.

23. Hitchcock, T. M., Dong, L., Connor, E. E., Meira, L. B., Samson, L. D., Wyatt, M. D., and Cao, W. (2004) *J Biol Chem* **279**, 38177-38183

24. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *Journal of Computational Chemistry* **4**, 187-217

25. MacKerell, A. D., and Banavali, N. K. (2000) *Journal of Computational Chemistry* **21**, 105-120

26. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *Journal of Physical Chemistry B* **102**, 3586-3616

27. O'Neill, R. J., Vorob'eva, O. V., Shahbakhti, H., Zmuda, E., Bhagwat, A. S., and Baldwin, G. S. (2003) *J Biol Chem* **278**, 20526-20532

28. Saparbaev, M., and Laval, J. (1998) *Proc Natl Acad Sci U S A* **95**, 8508-8513

29. Barrett, T. E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., and Pearl, L. H. (1998) *Cell* **92**, 117-129

30. Barrett, T. E., Scharer, O. D., Savva, R., Brown, T., Jiricny, J., Verdine, G. L., and Pearl, L. H. (1999) *EMBO J* **18**, 6599-6609

31. Dong, L., Mi, R., Glass, R. A., Barry, J. N., and Cao, W. (2008) *DNA Repair (Amst)* **7**, 1962-1972

32. Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995) *Nature* **373**, 487-493

33. Wibley, J. E., Waters, T. R., Haushalter, K., Verdine, G. L., and Pearl, L. H. (2003) *Mol Cell* **11**, 1647-1659

34. Drohat, A. C., and Stivers, J. T. (2000) *Biochemistry* **39**, 11865-11875

35. Gallinari, P., and Jiricny, J. (1996) *Nature* **383**, 735-738

36. Hang, B., Downing, G., Guliaev, A. B., and Singer, B. (2002) *Biochemistry* **41**, 2158-2165

37. Baker, D., Liu, P., Burdzy, A., and Sowers, L. C. (2002) *Chem Res Toxicol* **15**, 33-39

38. Saparbaev, M., Langouet, S., Privezentzev, C. V., Guengerich, F. P., Cai, H., Elder, R. H., and Laval, J. (2002) *J Biol Chem* **277**, 26987-26993

39. Lutsenko, E., and Bhagwat, A. S. (1999) *J Biol Chem* **274**, 31034-31038

40. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18**, 2714-2723

41. Liu, P., Theruvathu, J. A., Darwanto, A., Lao, V. V., Pascal, T., Goddard, W., 3rd, and Sowers, L. C. (2008) *J Biol Chem* **283**, 8829-8836

42. Suzuki, T., Matsumura, Y., Ide, H., Kanaori, K., Tajima, K., and Makino, K. (1997) *Biochemistry* **36**, 8013-8019

43. Kawase, Y., Iwai, S., Inoue, H., Miura, K., and Ohtsuka, E. (1986) *Nucleic Acids Res* **14**, 7727-7736

44. Parker, J. B., Bianchet, M. A., Krosky, D. J., Friedman, J. I., Amzel, L. M., and Stivers, J. T. (2007) *Nature* **449**, 433-437

45. Yang, W. (2006) *DNA Repair (Amst)* **5**, 654-666

46. Yang, W. (2008) *Cell Res* **18**, 184-197

Supplementary Content

Reagents, Media and Strains

All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO), Fisher Scientific (Suwanee, GA), or VWR (Suwanee, GA). Restriction enzymes, *Taq* DNA polymerase and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). BSA and dNTPs were purchased from Promega (Madison, WI). HiTrap chelating and Q columns were purchased from Amersham-Pharmacia Biotech (Piscataway, NJ). Oligodeoxyribonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA). LB medium was prepared according to standard recipes. Sonication buffer consisted of 20 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), 0.1 mM DTT (dithiothreitol), 0.15 mM PMSF (phenylmethylsulfonyl fluoride), and 50 mM NaCl. GeneScan stop buffer consisted of 80% formamide (Amresco, Solon, OH), 50 mM EDTA (pH 8.0), and 1% blue dextran (Sigma Chemicals). TB buffer (1 x) consisted of 89 mM Tris base and 89 mM boric acid. TE buffer consisted of 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA. *E. coli* host strain BH214 [*thr-1, ara-14, leuB6, tonA31, lacY1, tsx-78, galK2, galE2, dcm-6, hisG4, rpsL, xyl-5, mtl-1, thi-1, ung-1, tyrA*::Tn*10, mug*::Tn*10, supE44*, (DE3)] is a kind gift of Dr. Ashok Bhagwat (Wayne State University, Detroit, MI). *E. coli* strains BW1466 (CC106) and BW1739 (BW1466 but Δ*nfi*::(FRT-Spc-FRT) *nei*-1::cat *alkA*1) are kind gifts from Dr. Bernard Weiss at Emory University. *E. coli* strain JM109 [e14$^-$(McrA$^-$) *endA1, recA1, gyrA96, thi-1, hsdR17*(r_k^-, m_k^+), *supE44, relA1* Δ(*lac-proAB*), [F', *traD36, proAB, lacI*^{q}ZΔ*M15*]] is from our laboratory collection.

Plasmid Construction, Cloning, and Expression of *E. coli* MUG

The *E. coli* MUG gene was amplified by PCR using the forward primer Ec.MUG F (5'-TGG GGT ACC CCA TGG GTT GAG GAT ATT TTG GCT CCA GGG- 3'; the *Nco*I site is underlined) and the reverse primer Ec.MUG R (5'-CCC GGA TCC TTA TCG CCC ACG CAC TAC CAG CGC CTG GTC-3'; the *Bam*HI site is underlined). The PCR reaction mixture (50 µl) consisted of 8 ng of *E. coli* genomic DNA, 200 nM forward primer Ec.MUG F and reverse primer Ec.MUG R, 1 x Taq PCR buffer (New England Biolabs), 200 µM each dNTP, and 5 units of Taq DNA polymerase (New England Biolabs). The PCR procedure included a predenaturation step at 94 °C for 3 min; 30 cycles of three-step amplification with each cycle consisting of denaturation at 94 °C for 40 s, annealing at 60 °C for 40 s and extension at 72 °C for 1 min; and a final extension step at 72 °C for 10 min. The PCR product was purified with GeneClean 2 Kit (Qbiogene). Purified PCR product and plasmid pET32a(+) were digested with *Nco*I and *Bam*HI, purified with GeneClean 2 Kit and ligated according to the manufacturer's instructional manual. The ligation mixture was transformed into *E. coli* strain JM109 competent cells prepared by electroporation (1). The sequence of the *E. coli* MUG gene in the resulting plasmid (pET32a(+)-MUG) was confirmed by DNA sequencing.

To express the N-terminal His-6-tagged *E. coli* MUG, pET32a(+)-MUG was transformed into *E. coli* strain BH214 (*mug ung*) by electroporation (1). An overnight *E. coli* culture was diluted 100-fold into LB medium (500 ml) supplemented with 100 μg/ml ampicillin and grown at 37 °C with shaking at 250 rpm until the optical density at 600 nm reached about 0.6. After adding IPTG to a final concentration of 1 mM, the culture was grown at room temperature for an additional 16 h. Cells were collected by centrifugation at 4,000 rpm with GSA-10 rotor at 4 °C and washed once with pre-cooled sonication buffer.

To purify the *E. coli* MUG protein, the cell pellet from a 500-ml culture was suspended in a 10 ml sonication buffer and sonicated at output 5 for 3 x 1 min with a 5 min rest on ice between intervals. The sonicated solution was clarified by centrifugation at 12,000 rpm with SS-34 rotor at 4 °C for 20 min. The supernatant was transferred into a fresh tube and loaded onto a 1 ml HiTrap chelating column (GE Healthcare). The column was washed with chelating buffer A (20 mM sodium phosphate (pH 7.4), 500 mM NaCl and 2 mM imidazole). The bound protein in the column was eluted with a linear gradient of 0-100 % chelating buffer B (chelating buffer A and 500 mM imidazole).

Fractions of the eluate were analyzed by 12% SDS-PAGE and those fractions containing MUG (60% chelating buffer B) were pooled. The partially purified MUG protein was loaded onto a 1 ml HiTrap SP column, washed with HiTrap SP buffer A (20 mM HEPES-KOH (pH 8.0), 1 mM EDTA and 0.1 mM DTT) and eluted with a linear gradient of 0-100 % HiTrap SP buffer B (HiTrap SP buffer A and 1 M NaCl). Fractions containing MUG (30-50% HiTrap SP buffer B) were pooled and concentrated through Microcon YM 10 (Millipore). The protein concentration was quantified by the Bradford method using bovine serum albumin as a standard (2). The MUG protein was stored in aliquots at -80 °C. Prior to use, the protein was diluted in 1 x storage buffer (10 mM Tris-HCl (pH 8.0), 1 mM DTT, 1 mM EDTA, 200 ug/ml BSA, 50% Glycerol).

Site-Directed Mutagenesis

An overlapping extension PCR procedure was used for site-directed mutagenesis (3). Taking the construction of N18A as an example, the first round of PCR was carried out using pET32a(+)-MUG as template DNA with two pairs of primers, Ec.MUG F and N18A-R (5'-TGA AAG CCC AGG GGC GAT ACC GCA AAA CAC GAC CCG-3', the N18A site is underlined) pair & Ec.MUG R and N18A-F (5' GTG TTT TGC GGT ATC GCC CCT GGG CTT TCA TCC GCC-3', the N18A site is underlined) pair. The PCR mixtures (50 µl) contained 10 ng of pET32a(+)-MUG as a template, 200 nM of each primer pair, 200 µM each dNTP, 1 × *Taq* DNA polymerase buffer (New England Biolabs), and 5 units of *Taq* DNA polymerase (New England Biolabs). The PCR procedure included a predenaturation step at 95 °C for 3 min; 30 cycles of three-step amplification with each cycle consisting of denaturation at 95 °C for 50 s, annealing at 65 °C for 50 s and extension at 72 °C for 1 min; and a final extension step at 72 °C for 10 min. The resulting two expected PCR fragments were used for overlapping PCR to introduce the desired mutation. This second run of the PCR reaction mixture (100 µl),

which contained 1 µl of each of the first run PCR products, 100 µM each dNTP, 1 × *Taq* DNA polymerase buffer, and 5 units of *Taq* DNA polymerase, was initially carried out with a predenaturation at 95 °C for 2 min, five cycles with each cycle of denaturation at 95 °C for 30 s and annealing and extension at 60 °C for 4 min, and a final extension at 72 °C for 5 min. Afterward, 100 nM outside primers (Ec.MUG F and Ec.MUG R) were added to the above PCR reaction mixture to continue the overlapping PCR reaction under the same reaction condition with 25 additional cycles. The PCR product was cloned into pET32a(+) as described above. The recombinant plasmid (pET32a(+)-MUG-N18A), confirmed by DNA sequencing, was transformed into *E. coli* strain BH214 (*mug ung*) by electroporation. The *E. coli* MUG-N18A protein was expressed and purified as described above.

Molecular Modeling and Molecular Dynamics Simulations

Construction of the initial protein models

Using the Swiss-Pdb Viewer (SPDBV) program (4), the model of E. coli MUG bound to the decamer was generated by performing a structural alignment between the crystal structure of MUG and the crystal structure of the UDG/decamer complex. The UDG structure was then removed, leaving a structural model of MUG bound to DNA. Mutations of MUG were generated with the program MODELLER-9v4 (5). Using the "selection.mutate" command within MODELLER, atomic internal coordinates were used to build new cartesian coordinates for mutant residues (S23A, N140H and N140M). The mutated enzyme was then structurally aligned with the bound wt MUG (as described above).

Analysis of the protein models

Molecular dynamics simulations were performed on the bound MUG structures using the CHARMM 32b1 molecular mechanics software package (6), and the CHARMM 27 force field (7,8). The solvent molecules were represented with the explicit TIP3P water model. A solvent box was constructed that resulted in a minimum water layer of 10 Å between the solute and the boundary of the box, which yielded ~17700 water molecules. Fourteen sodium atoms were added for electrical neutrality. Periodic boundaries were used to maintain the density of the system and particle-mesh Ewald summation was applied to represent the long distance electrostatic interactions. The starting structures were gently minimized with the adopted basis Newton-Raphson (ABNR) module in CHARMM, for 100 steps to remove any unfavorable van der Waals clashes. The system was heated for 200 ps, from 200 K to 300 K in increments of 1 K. Using an integration timestep of 2 fs, a canonical ensemble was generated for 2 ns of production.

Construction of the initial DNA models

Base flipping potentials of mean force (PMF) were constructed based on the dodecamer sequence d(GTCAGNGCATGG), where N was the base to be flipped out of the helix. This sequence was chosen because it has been used in many base flipping studies previously and is amenable to validation (9-11). Using the program 3DNA (12),

an initial structure corresponding to the sequence d(GTCAGCGCATGG) was constructed. The base complementary to N was systematically mutated *in silico* to guanine, adenine, cytosine and thymine. The N in each of these DNA models was mutated from cytosine into uracil and xanthine. Mutations of the DNA were performed by first removing the base of interest from the DNA coordinate file (pdb file created with 3DNA). Coordinates for the new sequence, which included the mutated base, were then generated in CHARMM using the "ic build" command.

Generation of base-flipping potentials of mean force

Starting from these eight models of B-form DNA, umbrella sampling was performed to calculate the PMF (potential of mean force) associated with flipping the deaminated DNA bases following the previously described methods (9). A pseudodihedral angle, described previously (9), is used to define the base-flipping reaction coordinate. This pseudodihedral angle was defined through the centers of mass corresponding to a) the base pair 3' to the flipping base; b) the sugar 3' to the flipping base; c) the sugar of the flipped base; and d) the flipped base. This is the same pseudodihedral angle used as a flipping reaction coordinate in previous work (9). Previously, the PMF profiles for guanine flipping (G-flipping) and cytosine flipping (C-flipping) in an explicit solvent were both reported using the CHARMM 27 force field. The free energy difference for C-flipping was 15-18 kcal/mol, and for G-flipping 18-21 kcal/mol. The corresponding molecular dynamics simulations were run in an implicit generalized Born solvent. The PMF profiles were created by incrementing the pseudodihedral angle 5° in each simulation for 0°-360° (72 windows). A pseudodihedral angle of 0°-30° is defined as the base-paired state and an angle of 190° is defined as the base-opened or flipped out state, consistent with previous studies (9). Starting structures for these simulations were created by minimizing 100 steps with the adopted basis Newton-Raphson algorithm, and using the miscellaneous mean field potential (MMFP) module in the CHARMM package to increment the pseudodihedral angle with a force constant of 10,000 kcal/mol/rad^2. Starting structures were incremented ±5° from the final structures of the previous minimization. These starting structures were then used in the simulations. A harmonic umbrella potential $w_i(x) = k_i (x - x_i)^2$ was used to restrain the pseudodihedral angle with a force constant (k_i) of 1000 kcal/mol/rad^2. Positional restraints (force constant of 2 kcal/mol/rad^2) were applied to the four terminal bases to prevent fraying. The systems were heated for 60 ps from 250 K to 300 K, and equilibrated at 300 K. Langevin dynamics were used, with an integration timestep of 1 fs, to construct a canonical ensemble. The pseudodihedral values were recorded throughout the trajectory, and used to calculate a probability distribution. The weighted histogram analysis method (WHAM) was used to create unbiased PMF profiles (13). The implicitly solvated PMF profiles showed exaggerated free energy differences of ~30 kcal/mol for G-flipping, and ~31 kcal/mol for C-flipping. While quantitative differences are expected using implicit solvent models, similarities in the shape of the PMF profiles and the qualitative trend in free energies suggest this approach can provide useful information regarding base-flipping thermodynamics.

Description of Additional Mutants

An I17L substitution in motif 1 substantially reduced both the XDG and UDG activity (Fig. S2-A). Evidently, the UDG activity was reduced to such an extent that it is not detectable even when the enzyme was in excess (E:S = 10:1) (Fig. S2-A). Comparing I17L with the wt enzyme, the apparent rate constants for excision of xanthine by the mutant ranged from 10-fold lower for C/X, to 14-fold for G/X, to 16-fold for A/X, to 14-fold for T/X (Table 1). The single-stranded xanthine DNA glycosylase activity of I17L mutant was about 16-fold lower (Table 1). Given that the backbone amide of I17 in the wt enzyme forms a hydrogen bond with C^2-carbonyl oxygen of uracil (14), it is conceivable that amino acid substitutions at I17 could significantly impact DNA glycosylase activity.

G20 in MUG is equivalent to Y147 in hUNG (also called hUDG). Y147 in hUDG forms a rigid binding pocket to distinguish thymine from uracil (15,16). Y147 is completely conserved in family 1 UDGs but not in other families (Fig. S1). G20Y mutant was not active toward any deaminated base (data not shown), suggesting that the bulky Tyr residue was not tolerated in MUG/TDG family enzymes.

L21H mimicked family 1 UDG enzymes while A24M resembled the family 3 SMUG1 enzymes. S22M corresponded to equivalent residues in hTDG (Fig. S1). The activity profiles of these three mutants were similar to the wt enzyme with significant cleavage activity on X and U (data not shown), suggesting that these substitutions did not play a major role in determining the xanthine nor the uracil DNA glycosylase activities.

The mutational effect caused by L144S is similar to I17L, i.e., a significant reduction in XDG activity and a loss of UDG activity (Fig. S2-D and Table 1). It has been shown that L144 is part of a wedge that penetrates the DNA and interacts with the widowed guanine once the uracil base is flipped out (17). The backbone of L144 forms a hydrogen bond with the side chain of R146. It is proposed that MUG utilizes a push mechanism mediated by the P141-R146 loop to flip the uracil (or xanthine) out of the helix (17). The L144S mutation may affect the ability of this loop to stabilize the flipped conformation of the deaminated base and act as an effective wedge.

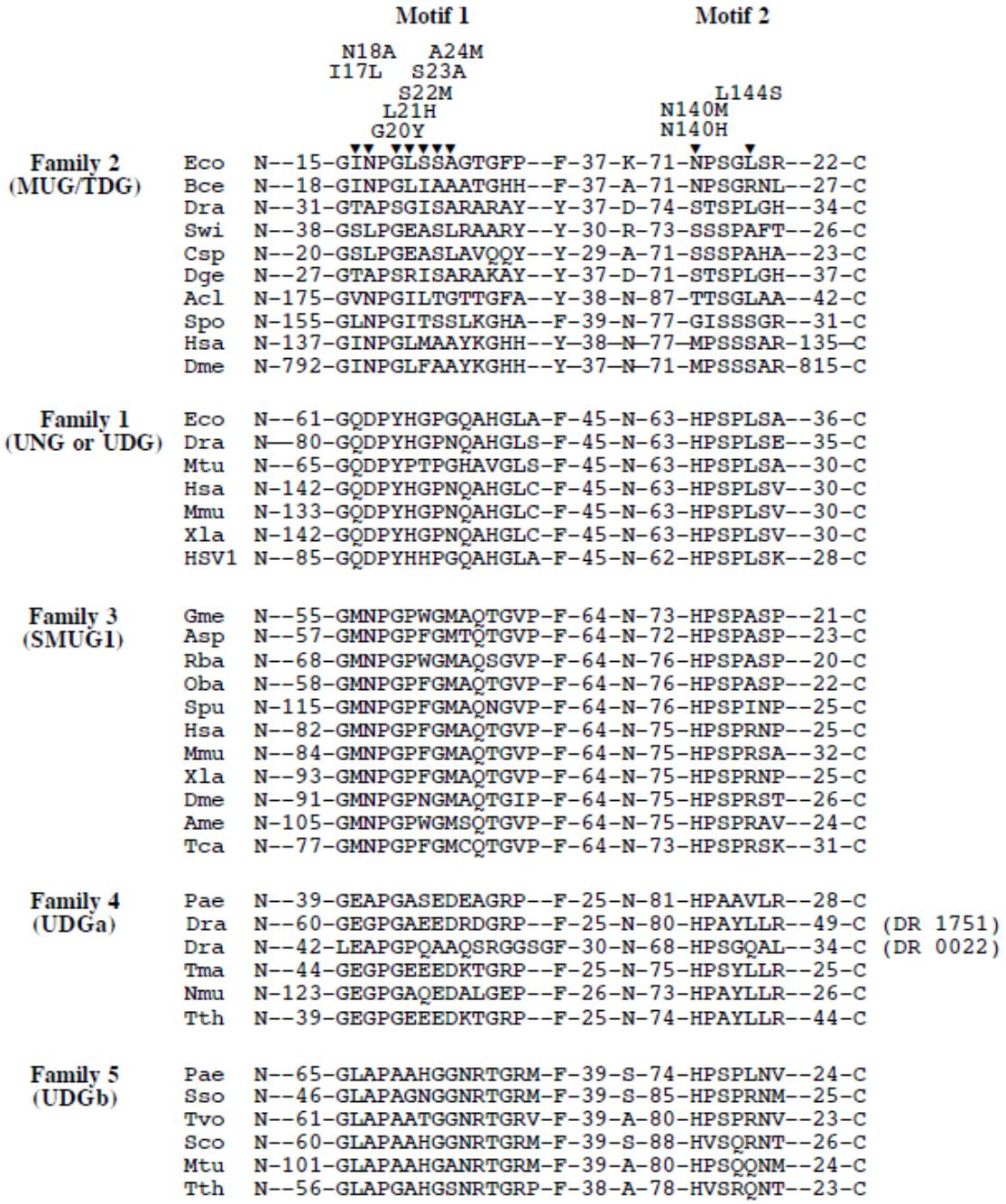

Figure S1. Sequence alignment of UDG superfamily. Genbank accession numbers are shown after the species names. **Family 2 (MUG/TDG):** Eco, *Escherichia coli*, P0A9H1; Bce, *Burkholderia cenocepacia* HI2424, YP_836419; Dra, *Deinococcus radiodurans* R1, NP_294438; Swi; *Sphingomonas wittichii* RW1, ZP_01607068; Csp, *Caulobacter sp.* K31, ZP_01418424.1; Dge, *Deinococcus geothermalis* DSM 11300, YP_605182.1; Acl, *Aspergillus clavatus* NRRL 1, XP_001268386.1; Spo, *Schizosaccharomyces pombe*, O59825; Hsa, *Homo sapiens*, NP_003202; Dme, *Drosophila melanogaster*, CAB93525; **Family 1 (UDG):** Eco, *Escherichia coli*, NP_289138; Dra, *Deinococcus radiodurans* R1, NP_294412; Mtu, *Mycobacterium*

tuberculosis H37Rv**,** CAB05436.1; Hsa, *Homo sapiens*, NP_003353; Mmu, *Mus musculus*, NP_035807; Xla, *Xenopus laevis*, NP_001085412; HSV1, *Herpes Simplex Virus-1*, 1UDI. **Family 3 (SMUG1):** Gme, *Geobacter metallireducens* GS-15, YP_383069; Asp, *Azoarcus sp.* BH72, YP_935478; Rba, *Rhodopirellula baltica* SH 1, NP_869403; Oba, *Opitutaceae bacterium* TAV2, ZP_02013615.1; Spu, *Strongylocentrotus purpuratus*, XP_782746.1; Hsa, *Homo sapiens*, NP_055126; Mmu, *Mus musculus*, NP_082161; Xla, *Xenopus laevis*, AAD17300; Dme, *Drosophila melanogastser*, NP_650609.1; Ame, *Apis mellifera*, XP_396883.2; Tca, *Tribolium castaneum*, XP_971699.1. **Family 4 (UDGa):** Pae, *Pyrobaculum aerophilum* str. IM2, NP_558739.1; Dra (DR 1751), *Deinococcus radiodurans* R1, NP_295474; Dra (DR 0022), *Deinococcus radiodurans* R1, AAF09614; Tma, *Thermotoga maritima* MSB8, NP_228321.1; Nmu, *Nitrosospira multiformis*, YP_412806; Tth, *Thermus thermophilus* HB27, YP_004341.1. **Family 5 (UDGb):** Pae, *Pyrobaculum aerophilum* str. IM2, NP_559226; Sso**,** *Sulfolobus solfataricus* P2, NP_344053.1; Tvo, *Thermoplasma volcanium* GSS1, NP_111346.1; Sco, *Streptomyces coelicolor* A3(2), NP_626251.1; Mtu, *Mycobacterium tuberculosis* H37Rv**,** P64785 (Rv1259); Tth, *Thermus thermophilus* HB27, YP_004757.1.

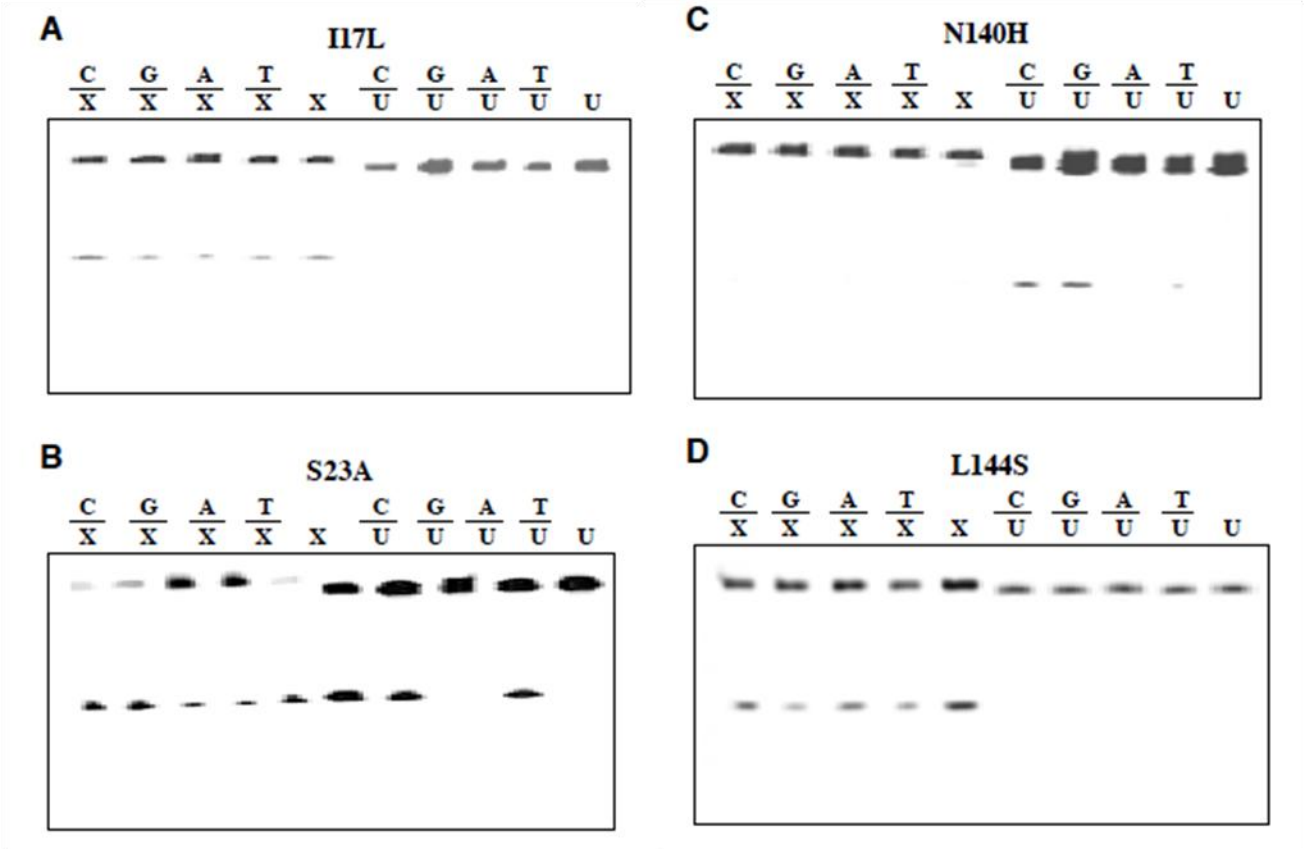

Figure S2. Glycosylase activity of I17L, S23A, N140H and L144S mutants of *E. coli* MUG on X-, U-containing substrates. Cleavage reactions were performed as described in Experimental Procedures with 100 nM *E. coli* MUG protein and 10 nM substrate. **A.** I17L. **B.** S23A. **C.** N140H. **D.** L144S.

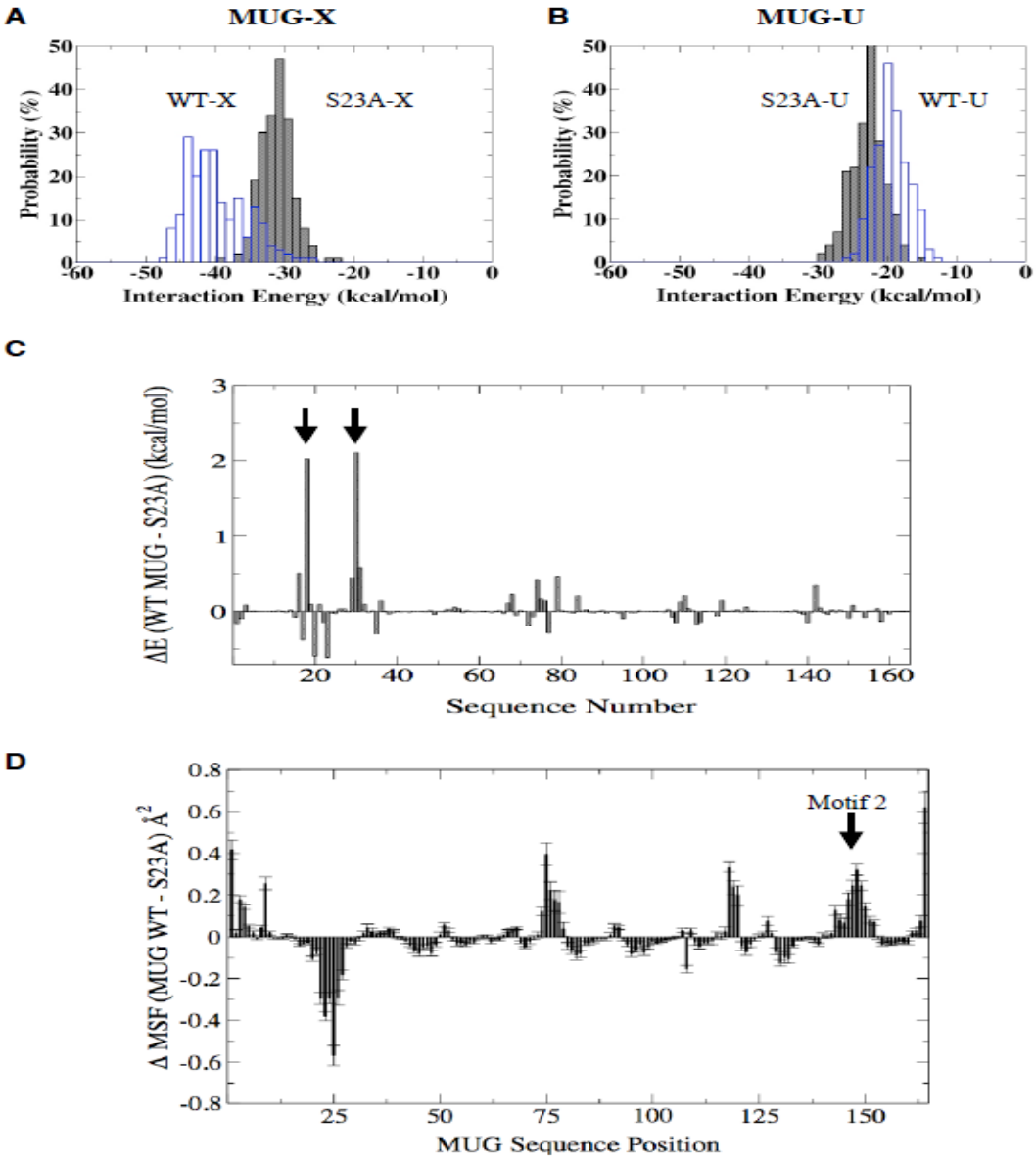

Figure S3. Interactions of S23A with xanthine and uracil. A. Energetics of *E. coli* MUG interactions with G/X. Blank bars, MUG-WT; solid bars, MUG-S23A. **B.** Energetics of *E. coli* MUG interactions with G/U. Blank bars, MUG-WT; solid bars, MUG-S23A. **C.** Differences between MUG-WT and MUG-S23A in energy of substrate (uracil) interactions per amino acid. Positive values indicate a more favorable interaction with MUG-S23A **D.** Difference in isotropic mean squared fluctuations between the MUG-WT and MUG-S23A. MSF values were calculated within CHARMM using the "coor dyna" command, and the error bars correspond to the standard error ($\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\Delta MSF_i - \langle \Delta MSF \rangle\right)^2}$) / $\sqrt{N}$) of the ΔMSF values over the MD trajectory. Positive MSF differences indicate that C-□'s in the S23A mutant are more rigid.

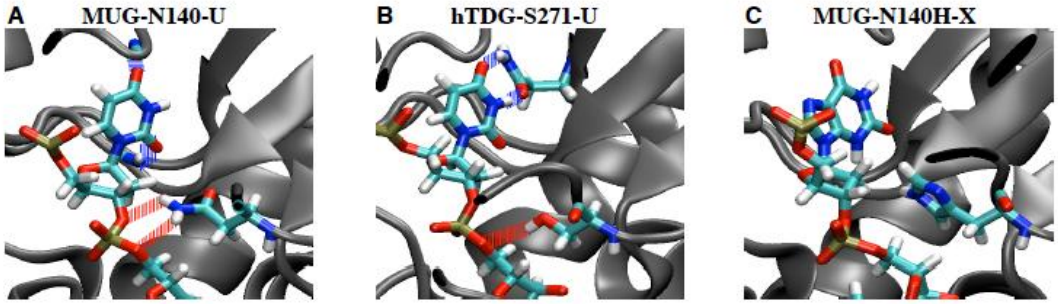

Figure S4. Modeled structures of *E. coli* MUG and human TDG. A. Interactions of the sidechain of N140 with 3'-phosphate in the DNA backbone in *E. coli* MUG. DNA and N140 are shown in color. **B.** Interactions of the sidechain of S271 with 3'-phosphate in the DNA backbone in human TDG. DNA and S271 are shown in color. **C.** Lack of interactions between *E. coli* MUG-N140H and xanthine. DNA and N140H are shown in color.

References

1. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1997) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc, New York
2. Coligan, J. E., Dunn, B. M., Ploegh, H. L., Speicher, D. W., and Wingfield, P. T. (1998) *Current Protocols in Protein Science*, John Wiley & Sons, Inc., New York
3. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* **77**, 51-59
4. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18**, 2714-2723
5. Sali, A. B., T. L. . (1993) *Journal of Molecular Biology* **234**, 779-815
6. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *Journal of Computational Chemistry* **4**, 187-217
7. MacKerell, A. D., and Banavali, N. K. (2000) *Journal of Computational Chemistry* **21**, 105-120
8. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *Journal of Physical Chemistry B* **102**, 3586-3616
9. Priyakumar, U. D., and MacKerell, A. D., Jr. (2006) *Chem Rev* **106**, 489-505
10. Priyakumar, U. D., and MacKerell, A. D. (2006) *Journal of Chemical Theory and Computation* **2**, 187-200
11. Varnai, P., and Lavery, R. (2002) *J Am Chem Soc* **124**, 7272-7273
12. Lu, X. J., and Olson, W. K. (2003) *Nucleic Acids Res* **31**, 5108-5121
13. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A., and Rosenberg, J. M. (1992) *Journal of Computational Chemistry* **13**, 1011-1021
14. Barrett, T. E., Scharer, O. D., Savva, R., Brown, T., Jiricny, J., Verdine, G. L., and Pearl, L. H. (1999) *EMBO J* **18**, 6599-6609

15. Kavli, B., Slupphaug, G., Mol, C. D., Arvai, A. S., Peterson, S. B., Tainer, J. A., and Krokan, H. E. (1996) *EMBO J* **15**, 3442-3447
16. Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E., and Tainer, J. A. (1995) *Cell* **80**, 869-878
17. Barrett, T. E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., and Pearl, L. H. (1998) *Cell* **92**, 117-129

CHAPTER FOUR

A NEW FAMILY OF DEAMINATED REPAIR ENZYMES IN URACIL DNA GLYCOSYLASE SUPERFAMILY

I. Abstract

Previous investigation identifies five families within the uracil DNA glycosylase (UDG) superfamily. All enzymes within the superfamily studied thus far contain uracil DNA glycosylase activity, although some members in certain families possess additional deaminated base repair activity. We identified a new class of DNA glycosylases in the UDG superfamily that did not show UDG activity, instead, they acted as hypoxanthine DNA glycosylase. Molecular modeling and structure-guided mutational analysis allowed us to identify a new catalytic center in this class of DNA glycosylases. Based on the biochemical properties and phylogenetic analysis, we propose this new class of DNA repair glycosylases as family 6 enzymes and designate it as HDG family.

II. Introduction

The uracil DNA glycosylase (UDG) superfamily consists of five families based on conserved motifs and structural similarity. Structurally, they are organized by a four-stranded β-sheet surrounded by α-helices, while functionally, all of the DNA glycosylases within the superfamily studied thus far are proven biochemically as uracil DNA glycosylase. Due to the need of removing deaminated cytosine from the genome, almost all organisms contain at least one uracil DNA glycosylase. Uracil-N-glycosylases (UNG) are family 1 UDG first discovered in *Escherichia coli* (1). All UNGs characterized show exquisite specificity towards uracil in both double-stranded (ds) and

single-stranded (ss) DNA (2). Family 2 enzymes are represented by human thymine DNA glycosylase (TDG), *E. coli* mismatch-specific uracil-DNA glycosylase (MUG), and a broad substrate specificity fission yeast *Schizosaccromyces pombe* TDG. This family exhibits more sequence and functional diversity. The family 3 SMUG1 enzymes (single-strand-selective monofuntional uracil-DNA glycosylase) are considered as a hybrid with a MUG/TDG-like motif 1 and UNG-like motif 2, which are active on both ds and ss U and XDG (3,4). While the prokaryotic family 4 UDG enzymes act on both single-stranded and double-stranded uracil-containing DNA, family 5 UDGs are found active towards G/U and to a lesser degree towards T/I substrates. Regardless which family an enzyme belongs to, a common biochemical feature is that they all have uracil DNA glycosylase activity.

In the course of studying deaminated DNA repair, we identified a new group of DNA glycosylases in the UDG superfamily. Unlike the previously known families, enzymes from this new family possess no uracil DNA glycosylase activity, and instead, exhibit repair activity towards hypoxanthine, a deamination product of adenine. The catalytic center is located to a completely conserved asparagine situated in a distinct orientation to the glycosidic bond. Based on the distinct repair capacity and active site architecture, we propose this class of repair enzymes as family 6 in the UDG superfamily. The discovery of this new family enzyme underlies the functional and catalytic diversity in the UDG superfamily and have implications in the evolution of repair enzymes.

III. Materials and Methods

A. Reagents, media and strains

All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO), Fisher Scientific (Suwanee, GA), or VWR (Suwanee, GA). Restriction enzymes, *Taq* DNA polymerase and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). BSA and dNTPs were purchased from Promega (Madison, WI). HiTrap chelating and Q columns were purchased from Amersham-Pharmacia Biotech (Piscataway, NJ). Oligodeoxyribonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA). LB medium was prepared according to standard recipes. Sonication buffer consisted of 50 mM HEPES-KOH (pH 7.4), 1 mM EDTA (pH 8.0), 2.5 mM DTT, and 0.15 mM PMSF, 10% glycerol and 50 mM NaCl. GeneScan stop buffer consisted of 80% formamide (Amresco, Solon, OH), 50 mM EDTA (pH 8.0), and 1% blue dextran (Sigma Chemicals). TE buffer consisted of 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA. *E. coli* host strain BH214 (*thr-1, ara-14, leuB6, tonA31, lacY1, tsx-78, galK2, galE2, dcm-6, hisG4, rpsL, xyl-5, mtl-1, thi-1, ung-1, tyrA*::Tn*10*, *mug*::Tn*10*, *supE44*, (DE3)) is a kind gift from Dr. Ashok Bhagwat (Wayne state university, Detroit, MI) and JM109 is from our laboratory collection. CC106, also called BW1466 (P90C (*ara*Δ(*proB-lac*)$_{xlll}$) *F' lacI378 lacZ proB*$^{+}$) and BW1506 (CC106 *nfi-1::cat*) are a kind gift of Dr. Bernard Weiss (Emory University, Atlanta, GA) (5,6). Genomic DNA from *Methanosarcina barkeri strain Fusaro* and *Methanosarcina acetivorans* C2A are a kind gift from Dr. William Metcalf (University of Illinois, Urbana, IL).

B. Cloning, Expression and Purification of Mba DNA Glycosylase

The *Methanosarcina barkeri* DNA glycosylase gene (GenBank accession number: YP304295.1) was amplified by PCR using the forward primer Mba.F (5'- GGG AAT TCC ATA TGA AAA AAC AAG GTT TCC CAC CAG TCA TT -3'; the NdeI site is underlined) and the reverse primer Mba.R (5'- CCG CTC GAG ACG CTT GAA AGC TTC TGC CCA TCC CGA TTT -3'; the XhoI site is underlined). The PCR reaction mixture (50 µL) consisted of 8 ng of Mba genomic DNA, 200 nM forward primer and reverse primer, 1 x Taq PCR buffer (New England Biolabs), 200 µM each dNTP, and 5 units of Taq DNA polymerase (New England Biolabs). The PCR procedure included a predenaturation step at 94°C for 3 min, 30 cycles of three-step amplification with each cycle consisting of denaturation at 94°C for 40 sec, annealing at 60°C for 40 sec and extension at 72°C for 1 min, and a final extension step at 72°C for 10 min. The PCR product was purified with Gene Clean 2 Kit (Qbiogene). Purified PCR product and plasmid pET21a were digested with NdeI and XhoI, and followed by purifying gene fragments with Gene Clean 2 Kit and ligated according to the manufacturer's instruction manual. The ligation mixture was transformed into *E. coli* strain JM109 competent cells prepared by electroporation (7). The sequence of the Mba DNA glycosylase gene in the resulting plasmid (pET21a-Mba) was confirmed by DNA sequencing.

To express the C-terminal His-6-tagged Mba glycosylase, the plasmid pET21a-Mba was transformed into *E. coli* strain BH214 *(ung*$^-$ *mug*$^-$) by electroporation (7). An overnight *E. coli* culture was diluted 100-fold into LB medium supplemented with 100 *µ*g/mL ampicillin. The *E. coli* cells were grown at 37°C while being shaken at 250 rpm until the optical density at 600 nm reached approximately 0.6. After adding IPTG to a

final concentration of 1 mM, the culture was grown at room temperature for an additional 16 h. The cells were collected by centrifugation at 4,000 rpm at 4°C and washed once with pre-cooled sonication buffer.

To purify the Mba DNA glycosylase protein, bacterial cells from a 500 ml culture grown to late exponential phase were harvested by centrifugation at 4,000 rpm for 10 min. The cell pellet was suspended in 7 ml of lysis buffer (20 mM Tris-HCl (pH 7.5), 1 mM EDTA (8.0), 0.1 mM DTT, 0.15 mM PMSF and 50 mM NaCl) and followed by sonication at output 5 for 3 x 1 min with 5 min rest on ice between intervals using a Sonifier Cell Disruptor 350 (Branson). The lysate was clarified by centrifugation at 12,000 rpm for 20 min and filtered through a 25 mm GD/X syringe filter (Whatman). The supernatant was transferred into a fresh tube and loaded onto a 1 ml Hi-Trap chelating column (GE Healthcare). The column was washed with 5 ml of chelating buffer A (20 mM sodium phosphate (pH 7.4), 500 mM NaCl and 2 mM imidazole). The bound protein in the column was eluted with a linear gradient of 0-100 % chelating buffer B (chelating buffer A and 500 mM Imidazole).

Fractions of the eluate were analyzed by 12% SDS-PAGE and those fractions containing Mba glycosylase protein (40% chelating buffer B) were pooled. The partially purified protein was then loaded onto a HiTrap SP column, washed with 5 ml of HiTrap SP buffer A (20 mM HEPES (pH 8.0), 1 mM EDTA and 0.1 mM DTT) and eluted with a linear gradient of 0-100 % HiTrap SP buffer B (HiTrap SP buffer A and 1 M NaCl). Fractions containing Mba glycosylase (30-50% HiTrap SP buffer B) were pooled and concentrated through Microcon YM 10 (Millipore). The protein concentration was

quantified by SDS-PAGE analysis using bovine serum albumin as a standard. The Mba glycosylase protein was stored in aliquotes at -80°C. Prior to use, the protein was diluted in equal volume of 2 x storage buffer (20 mM Tris-HCl (pH 8.0), 2 mM DTT, 2 mM EDTA, 400 ug/ml BSA, 100% Glycerol).

C. Cloning, Expression and Purification of Mac DNA Glycosylase

The *Methanosarcina acetivorans* DNA glycosylase gene (GenBank accession number: YP615428.1) was amplified by PCR using the forward primer Mac.F (5'-GGGAATTCCATATGATAAAGCGAGGTTTTCCTGCAGTCCTT-3'; the NdeI site is underlined) and the reverse primer Mac.R (5'- CCGCTCGAGATGCCTGAAAACAG CCTCCCACTCCGATTC-3'; the XhoI site is underlined). The cloning, expression and purification steps were the same as described above.

D. Site-Directed Mutagenesis

Site-directed mutagenesis of Mba glycosylase was performed using an overlapping extension PCR procedure (8). The first round of PCR was carried out using pET21a-Mba as template DNA with two pairs of primers, Mba.F and N39A-R (5'-CCTCCAGAAATCAGCGCCTGGATGCCCATAGTATTG-3') pair and and N39A-F (5'-TATGGGCATCCAGGCGCTGATTTCTGGAGGCTGCTT-3') and Mba.R pair; Mba.F and N39D-R (5'-CCTCCAGAAATCATCGCCTGGATGCCCATA GTATTG-3') pair and N39D-F (5'-TATGGGCATCCAGGCGATGATTTCTGGAGGC TGCTT-3') and Mba.R pair; Mba.F and N39Q-R (5'-CCTCCAGAAATCTTGGCCTG GATGCCCATAGTATTG-3') pair, and N39Q-F (5'-TATGGGCATCCAGGCCAAGA TTTCTGGAGGCTGCTT-3') and Mba.R pair, respectively. The PCR mixtures (50 µl)

contained 10 ng of pET21a-Mba as a template, 200 nM each primer, 200 µM each dNTP, 1× *Taq* DNA polymerase buffer, and 5 unit of *Taq* DNA polymerase. The PCR procedure included a predenaturation step at 95°C for 3 min, 30 cycles of three-step amplification with each cycle consisting of denaturation at 95°C for 50 s, annealing at 65°C for 50 s and extension at 72°C for 1 min, and a final extension step at 72°C for 10 min. The resulting two expected DNA fragments were used for overlapping PCR to introduce the desired mutation. This second run of PCR reaction mixture (100 µl), which contained 1 µl of each of the first run PCR products, 100 µM each dNTP, 1× *Taq* DNA polymerase buffer, and 5 units of *Taq* DNA polymerase, was initially carried out with a predenaturation at 95°C for 2 min, five cycles with each cycle of denaturation at 95°C for 30 s and annealing and extension at 60°C for 4 min, and a final extension at 72°C for 5 min. Afterward, 100 nM outside primers (Mba.F and Mba.R) were added to the above PCR reaction mixture to continue the overlapping PCR reaction under the same reaction condition with 25 additional cycles. The PCR product was cloned into pET21a as described above. The recombinant plasmids (pET21a-N39A pET21a-N39D, pET21a-N39Q) containing the desired mutations were confirmed by DNA sequencing and transformed into *E. coli* strain BH214 by electroporation. The Mba glycosylase mutant proteins were expressed and purified similarly as the w Mba glycosylase as described above.

E. Oligodeoxynucleotide Substrates

The fluorescently labeled inosine- and uridine-containing oligonucleotides were ordered from IDT, purified by PAGE, and dissolved in TE buffer at a final concentration

of 10 μM. The two complementary strands with the unlabeled strand in 1.2-fold molar excess were mixed, incubated at 85°C for 5 min, and allowed to form duplex DNA substrates at room temperature for more than 30 min. The xanthosine- and oxanosine-containing oliogonucleotide were prepared as previously described (9,10).

F. DNA Glycosylase Activity Assay

DNA glycosylase cleavage assays were performed at 37°C for 60 min in a 10 µl reaction mixture containing 10 nM oligonucleotide substrate, 100 nM glycosylase protein, 20 mM Tris-HCl (pH 7.5), 5 mM EDTA and 2 mM 2-mercaptoethanol. The resulting abasic sites were cleaved by incubation at 95°C for 5 min after adding 1 µl of 1 N NaOH. Reactions were quenched by addition of an equal volume of GeneScan stop buffer. Samples (4 µl) were loaded onto a 7 M urea-10% denaturing polyacrylamide GeneScan gel (acrylamide : bisacrylamide = 19:1, 1 x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA)). Electrophoresis was conducted at 1500 V for 1.5 h using an ABI 377 sequencer (Applied Biosystems). Cleavage products and remaining substrates were quantified using GeneScan analysis software.

G. Reversion of lacZ gene in E. coli CC106

The plasmids pET21a-Mba and pET21a-Mba-N39A were digested with EcoRI and XhoI and the fragment containing the Mba DNA glycosylase gene was cloned to pBluescript SK(+) to generate pBS-Mba and pBS-Mba-N39A. The resulting plasmids were transformed to E. coli strains CC106 and BW1506 (CC106 *nfi-1::cat*). Tester cultures, inoculated as a single colony, were grown in Luria-Bertani (LB) medium at 37°C for 16 hr. Overnight cultures (1 ml) were transferred to 4 ml fresh LB medium and

incubated at 37°C for 4-5 hr until OD_{600} reached 0.6. After adding IPTG to a final concentration of 1 mM, the cultures were incubated at 37°C for an additional 5 hr. *E. coli* cells (2×10^9) were incubated with 30 µl of 1 M $NaNO_2$ (30 mM final concentration) and 970 µl of 100 mM sodium acetate buffer (pH 4.6) at 37°C for 30 min. Top agar consisting of 0.5 % NaCl, 0.6% agar (Difco) and 0.2 mg/ml nutrient broth (Difco) was autoclaved and then maintained at 45°C in a water bath. After centrifugation, the treated cells were suspended with 1 ml of 10 mM $MgSO_4$. Molten top agar prepared above (8 ml) was added to the suspended cells and the mixtures were immediately overlaid onto a minimum lactose (ML) plate prepared according to the recipe as previously described (11). Cells were incubated at 37°C for four days to allow a few times of cell division to fix mutations in the presence of nutrient broth (12).

H. Phylogenetic Analysis

The phylogenetic tree was generated with neighbor-joining algorithm of the MEGA 5 software (http://www.megasoftware.net/) applied to a multiple alignment produced with the CLUSTAL W program. The parameters for pairwise alignment are: gap opening penalty, 10; gap extension penalty 0.1. The parameters for multiple alignment are: gap opening penalty, 10; gap extension penalty 0.2. Other parameters are: protein weight matrix, Blosum; residue-specific penalities: on; hydrophilic penalities: on; gap separation distance: 4; end gap separation: off.

I. Molecular Modeling

Sequence alignment was performed using CLUSTAL W. pairwise alignment of the amino acid sequence from the Mba glycosylase (YP_304295.1) and chain A of

2rba.pdb (human TDG) resulted in 19% identity between the two sequences. Based on these sequence alignments and the 2rba pdb structure, a homology model was constructed for Mba enzyme using the NEST program (13).

IV. Results and Discussion

A. Deaminated Base Repair Activities

Family 2 enzymes (MUG/TDG) in the UDG superfamily were previously found only in bacteria and eukaryotic systems. In a search (Psi BLAST) for uracil DNA repair enzymes in archaea, we found a hypothetic protein (YP_502486.1) in archaea *Methanospirillum hungatei* that only showed insignificant E-value (E = 0.41) with *E. coli* MUG.

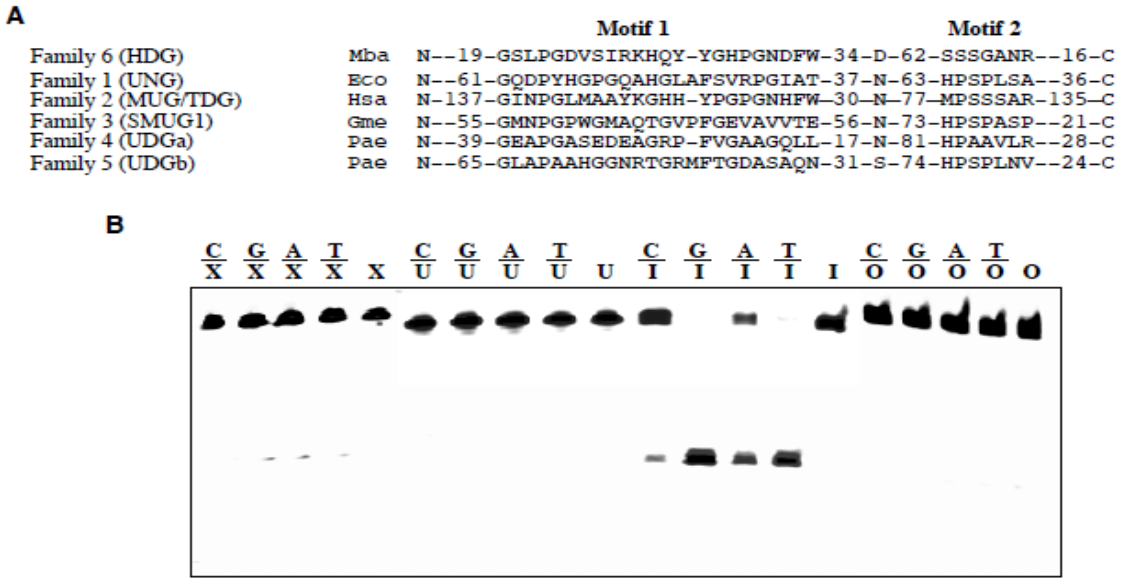

Figure 4.1 Sequence motifs and deaminated base repair activity from Mba DNA glycosylase. A. Sequence motifs of UDG superfamily. Genbank accession numbers are shown after the species names. Family 6 (HDG): Mba: YP_304295.1 Family 1 (UNG): Eco, Escherichia coli, NP_289138. Family 2 (MUG/TDG): Hsa, Homo sapiens, NP_003202. Family 3 (SMUG1): Gme, Geobacter metallireducens GS-15,YP_383069; Family 4 (UDGa): Pae, Pyrobaculum aerophilum str. IM2, NP_558739.1. Family 5 (UDGb): Pae, Pyrobaculum aerophilum str. IM2, NP_559226. B. Deaminated base repair activity from Mba DNA glycosylase. DNA glycosylase activity assays were performed as described in Materials and Methods with 100 nM wt Mba glycosylase protein and 10 nM substrate.

However, querying the GenBank with this protein led to identification of a large number of genes in archaea, eubacteria, and eukaryotes with significant homolog to UDG superfamily enzymes. Alignment of proteins from this family to the UDG family indicated significant sequence divergence in motifs I and II (Fig. 4.1A). Most notably, the position equivalent to the catalytic residue in *E. coli* UNG (family 1, D64) and *E. coli* MUG (family 2, N18) is occupied by a leucine residue. To investigate the repair activities of this class of enzymes, we expressed the homologous gene from *Methanosarcina barkeri* (Mba) in *E. coli* strain lacking both *ung* and *mug* and purified the protein to homologeneity. To our surprise, the recombinant protein completely lacked any detectable uracil DNA glycosylase activity on all four uracil-containing base pairs and the single-stranded uracil-containing DNA (Fig. 4.1B).

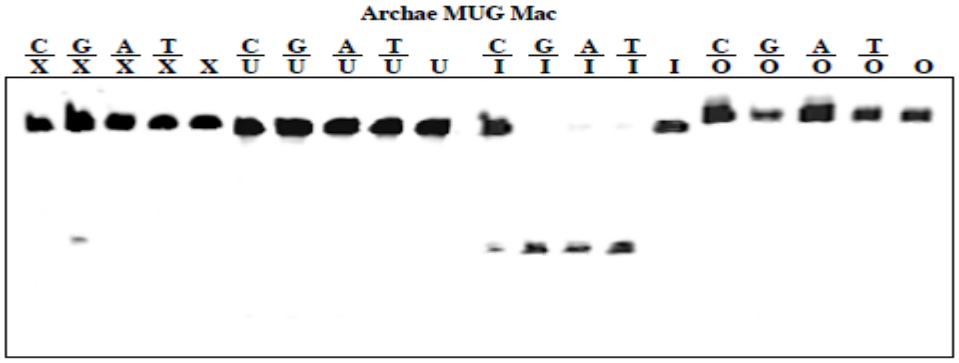

Figure 4.2 Deaminated base repair activity from Mac DNA glycosylase. DNA glycosylase activity assays were performed as described in Materials and Methods with 100 nM wt Mac glycosylase protein and 10 nM substrate.

On the other hand, hypoxanthine DNA glycosylase activity was found in all four double-stranded substrates (Fig. 4.1B). Some minor xanthine DNA glycosylase activities

were also detected in this Mba enzyme but no oxanine DNA glycosylase activities were detected (Fig. 4.1B).

To verify that the lack of UDG activity and existence of HDG activity is not unique to the Mba enzyme, we also investigated the deaminated DNA glycosylase activity in the homologous gene from *Methanosarcina acetivorans* (Mac). A similar glycosylase pattern was observed, indicating that this is a common repair property of this class of enzymes (Fig. 4.2).

B. Kinetic Analysis and Catalytic Center

Kinetic analysis showed that this glycosylase was most active on the G/I base pair with an apparent rate constant of 0.085 per min, followed by T/I, A/I and C/I base pairs (Table 4.1). No activity was detected on single-stranded inosine-containing DNA (Fig. 4.1B). The catalytic mechanisms in families 1 and 2 enzymes have been extensively studied. D64 in *E. coli* UNG has been identified as the key catalytic residue that activates a water molecule (14). In *E. coli* MUG, N18 utilizes the mainchain and sidechain oxygen to activate the water and mainchain amino to interact with uracil (15). However, the residue (Asp or Asn) that can perform the catalytic function is notably missing in the Mba enzyme. Instead, the equivalent position is occupied by a hydrophobic residue leucine (Fig. 4.1A). We constructed a model of the Mba enzyme based on the crystal structure of human TDG because they share some sequence homolog (Fig. 4.3A). Within the 10 Å radius of the AP site, the following potential catalytic residues were identified, N39 (7.08 Å), N84 (9.89 Å), N113 (8.62 Å), D74 (5.55 Å), D25 (9.46 Å), D86 (8.60 Å), E82 (3.05 Å), E85 (8.28 Å).

Table 4.1 Apparent rate constants of hypoxanthine DNA glycosylase activity in Mba DNA glycosylase (min–1)[a]

	Bottom Strand	Top Strand				Single Strand
		C	G	A	T	
WT	I	0.0092	0.085	0.014	0.026	n.a.[b]
D25I		0.0052	0.034	0.0099	0.024	n.a.
N39A		n.a.	n.a.	n.a.	n.a.	n.a.
N39D		n.a.	0.0011	n.a.	n.a.	n.a.
N39Q		n.a.	0.012	n.a.	n.d.[c]	n.a.
D74A		n.a.	0.0029	n.a.	n.a.	n.a.
D74N		n.a.	0.0018	n.a.	n.a.	n.a.
D86A		n.a.	0.011	n.a.	n.a.	n.a.
N113A		n.a.	0.061	n.a.	n.a.	n.a.
N141A		n.a.	0.093	n.a.	0.029	n.a.

[a]: The reactions were performed as described in Materials and Methods with 100 nM Mba MUG and 10 nM substrate, unless otherwise stated. The apparent rate constants were determined by fitting the time course data into a first-order rate equation using Deltagraph (SPSS Inc.). Data are an average of at least two independent experiments.
[b]: n.a.: No activity was detected under assay conditions.
[c]: n.d.: Not determined due to very low level of activity.

Among the eight residues identified, only N39 is completely conserved in the Mba enzyme and its homologs. D74, D86 and N113 are highly conserved. To assess the role of these residues in catalysis, D74, D86 and N113 were substituted by either Ala or Asn. Significant portion of HDG activity on the G/I substrate was retained, suggesting that

D74, D86 and N113 did not play a key role in catalysis (Table 4.1). Given the completely conserved nature, N39 was substituted with Ala, Asp and Gln.

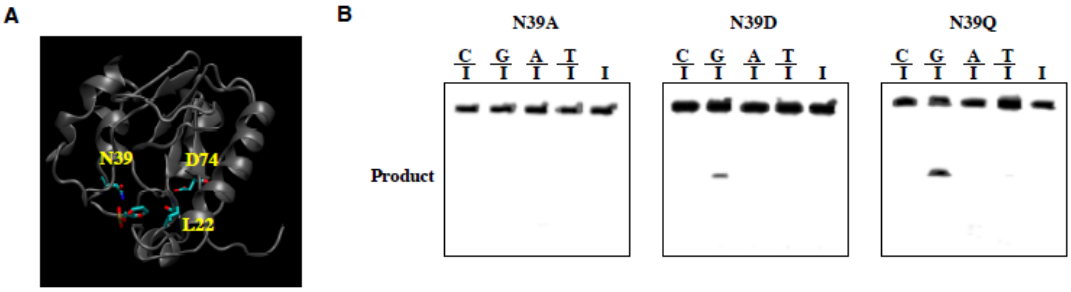

Figure 4.3 DNA glycosylase activity of N39 mutants. A. modeled structure of Mba DNA glycosylase. B. Hypoxanthine DNA glycosylase activity of N39A, N39D and N39Q mutants. DNA glycosylase activity assays were performed as described in Materials and Methods with 100 nM wt Mba glycosylase protein and 10 nM substrate.

The Ala substitution, which removed the functional group of the N39 sidechain, completely abolished the glycosylase activity (Fig. 4.3 and Table 4.1). Conversion of the sidechain from an amide group to a carboxyl group (N39D) rendered the enzyme active only on the G/I substrate (Fig. 4.3 and Table 4.1). Addition of a methylene group into the sidechain (N39Q) allowed the enzyme to retain significant activity on the G/I substrate but negligible activity on the T/I substrate (Fig. 4.3 and Table 4.1). These results indicate that N39 played a key role as the catalytic residue and the catalytic environment required the presence of an Asn residue in the catalytic center.

C. *In vivo* repair

To assess the role of the new group of DNA glycosylase may play *in vivo*, we adopted the lacZ-based genetic assay (5). *E. coli* CC106 strain contains G to A transition mutation at codon 461 that allows for detection of A to G revertants, which has been used to provide

evidence to implicate bacterial endonuclease V (encoded by *nfi* gene) as an

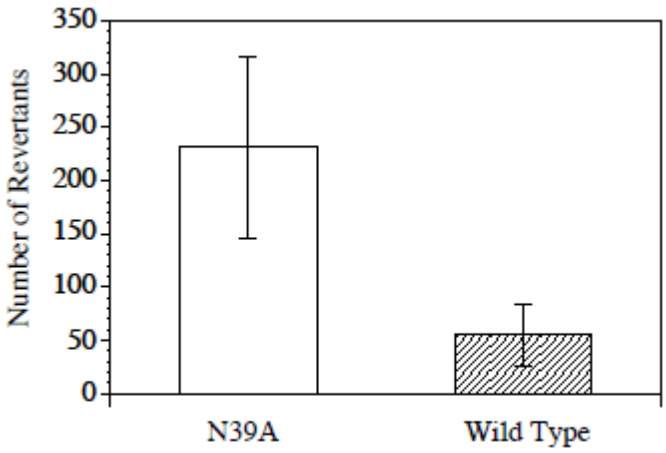

Figure 4.4 Role of Mba DNA glycosylase in vivo. Numbers of revertants were scored on minimal lactose plates after incubation at 37°C for four days. The data are the means ± SD of at least three independent experiments. Open bar, N39A mutant of Mba DNA glycosylase; hatched bar: wt Mba DNA glycosylase.

enzyme for the repair of adenosine deamination (6). It is known that DNA polymerases predominantly incorporate dCMP to pair with inosine in DNA templates (16). To access the role that the hypoxanthine DNA glycosylase activity may play *in vivo*, both a plasmid containing the wt Mba glycoysylase and a plasmid containing the N39A mutant were transformed into an *nfi* deficient CC106 strain. While the N39A mutant failed to suppress A to G mutations in the CC106 *nfi*$^-$ cells under nitrosative stress, the wt Mba DNA glycosylase reduced the number of revertants by more than four-fold (Fig. 4.4). These results suggest that the hypoxanthine DNA glycosylase activity in the Mba repair enzyme is responsible for the removal of hypoxanthine *in vivo*.

D. UDG-less Enzymes in UDG Superfamily

Since the discovery of the first uracil DNA glycosylase in *E. coli* in 1974, all the uracil DNA glycosylases studied possess UDG activity. While repair activities on other deaminated bases have been found in different families in the UDG superfamily, UDG activity is always present. This work for the first time reports a class of UDG enzymes without UDG activity. Instead, they exhibit significant repair activity on hypoxanthine, the deamination product of adenine. Data from the *in vivo* reversion assay are consistent with the notion that this class of enzymes acts as a hypoxanthine DNA glycosylase (Fig. 4.4). These observations suggest that different UDG families have evolved to possess distinct repair specificities during evolution. Similar to the family 1 UNG enzymes, the repair activity of this class of enzymes appears to be limited to hypoxanthine, indicating that it may have a narrow specificity. The HDG activity follows the order of G/I > T/I > A/I > C/I, which is consistent with the stability of inosine-containing base pairs (17-19). Therefore, the tendency of spontaneous base flipping appears to play an important role in determining the repair activity of this class of enzymes. Hypoxanthine differs from uracil by missing a C^2-keto and having an imidazole ring. It remains to be answered on how this class of enzyme specifically recognizes a hypoxanthine but not uracil. More structural studies are needed to address this essential question.

E. Unique Catalytic Center

The catalytic mechanisms in families 1 and 2 have been extensively studied. In family 1 UNG enzymes, a completely conserved Asp residue (D145 in human UNG) rotates 120° once bound to a uracil-containing DNA and acts as a general base to activate a bound water molecule (20-22). In family 2 enzymes, the corresponding position is

occupied by an Asn residue, which is proposed to activate a bound water molecule through sidechain and mainchain interactions (15,23). In the Mba enzyme, the equivalent position is occupied by a Leu residue (Fig. 4.1A). In the modeled structure, the L22 is not positioned toward the scissile bond (Fig. 4.3A). Among residues in the vicinity of the scissile bond, N39 emerged as the key catalytic residue that may perform the catalytic function similar to the Asp in family 1 and Asn in family 2. The evidence supports this notion includes, 1) N39 is completely conserved among the new class of enzymes; 2) N35 is better positioned in the modeled structure to activate a water molecule to potentiate an in-line attack on the glycosidic bond; 3) N39 mutant completely losses its catalytic activity.

F. A New Family Enzymes in UDG Superfamily

Based on the unprecedented lack of uracil DNA glycosylase activity, presence of HDG activity, and unique location of catalytic center, we performed a phylogenetic analysis within the UDG superfamily. As shown in Fig. 4.5, this new class of enzymes emerged as a distinct group within the superfamily. Therefore, we propose this class of enzymes as family 6 in the UDG superfamily and designate it as HDG family. The information provided for the family 6 enzymes could help us understand the multiplicity of UDG superfamily enzymes in a variety of organisms. For example, while a family 4 enzyme in *M. barkeri* may carry out the repair function for cytosine deamination damage by acting as a UDG, the family 6 enzyme described here could perform the necessary repair for adenine deaminaiton damage by acting as a HDG. How the UDG superfamily enzymes lost or acquired different deaminated base repair activity and relocated the catalytic

center is a subject of study that can reveal the adaptation of repair enzymes during millions of years of evolution.

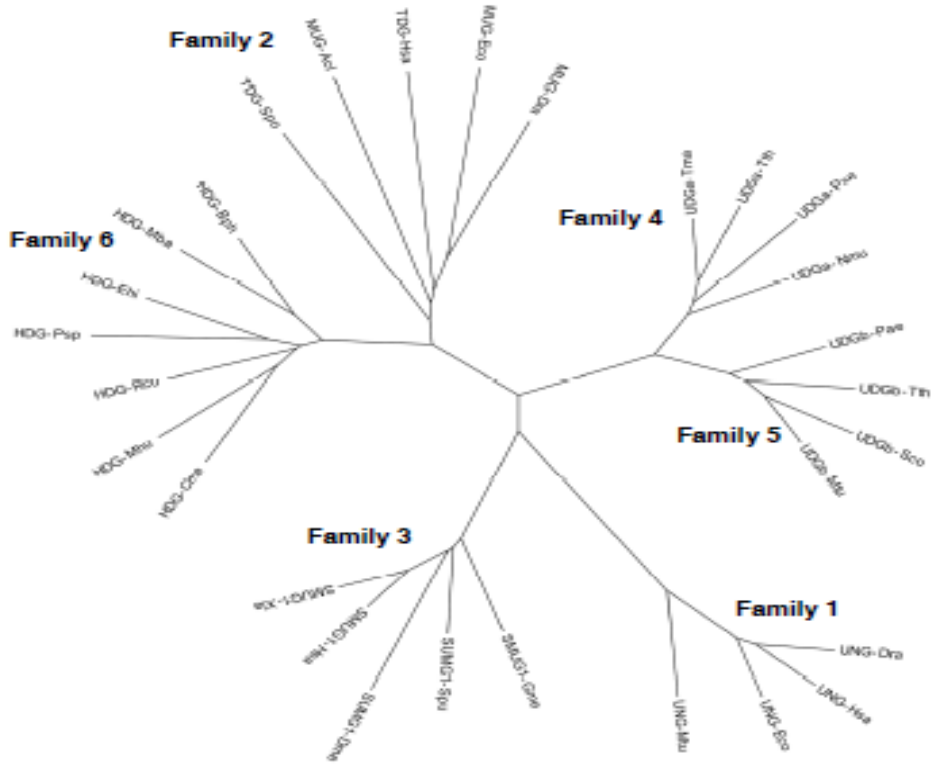

Figure 4.5. **Phylogenetic analysis of UDG superfamily.** The phylogenetic analysis was performed using the Neighbor-Joining method. The optimal tree with the sum of branch length = 14.40324302 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The analysis involved 29 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 118 positions in the final dataset. Evolutionary analyses were conducted in MEGA 5 (9).

V. Acknowledgements

This project was supported in part by CSREES/USDA (SC-1700274, technical contribution No. XXXX), Department of Defense-Army Research Office (W911NF-05-1-0335 and W911NF-07-1-0141).

VI. References

1. Lindahl, T. (1974) An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues. *Proc Natl Acad Sci U S A*, **71**, 3649-3653.
2. Parikh, S.S., Putnam, C.D. and Tainer, J.A. (2000) Lessons learned from structural results on uracil-DNA glycosylase. *Mutat Res*, **460**, 183-199.
3. Haushalter, K.A., Todd Stukenberg, M.W., Kirschner, M.W. and Verdine, G.L. (1999) Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. *Curr Biol*, **9**, 174-185.
4. Mi, R., Dong, L., Kaulgud, T., Hackett, K.W., Dominy, B.N. and Cao, W. (2009) Insights from xanthine and uracil DNA glycosylase activities of bacterial and human SMUG1: switching SMUG1 to UDG. *J Mol Biol*, **385**, 761-778.
5. Cupples, C.G. and Miller, J.H. (1989) A set of lacZ mutations in Escherichia coli that allow rapid detection of each of the six base substitutions. *Proc Natl Acad Sci U S A*, **86**, 5345-5349.
6. Schouten, K.A. and Weiss, B. (1999) Endonuclease V protects Escherichia coli against specific mutations caused by nitrous acid. *Mutat Res*, **435**, 245-254.
7. Sambrook, J. and Russell, D.W. (2001) *Molecular Cloning-A Laboratory Manual*. 3rd ed. Cold Spring Harbor laboratory Press, Cold Spring Harbor, New York.
8. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymease chain reaction. *Gene*, **77**, 51-59.

9. Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*, **24**, 1596-1599.

10. Feng, H., Klutz, A.M. and Cao, W. (2005) Active site plasticity of endonuclease V from Salmonella typhimurium. *Biochemistry*, **44**, 675-683.

11. Josephy, P.D. (2000) The Escherichia coli lacZ reversion mutagenicity assay. *Mutat Res*, **455**, 71-80.

12. Kim, S.R., Maenhaut-Michel, G., Yamada, M., Yamamoto, Y., Matsui, K., Sofuni, T., Nohmi, T. and Ohmori, H. (1997) Multiple pathways for SOS-induced mutagenesis in Escherichia coli: an overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc Natl Acad Sci U S A*, **94**, 13792-13797.

13. Petrey, D., Xiang, Z., Tang, C.L., Xie, L., Gimpelev, M., Mitros, T., Soto, C.S., Goldsmith-Fischman, S., Kernytsky, A., Schlessinger, A. *et al.* (2003) Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. *Proteins*, **53 Suppl 6**, 430-435.

14. Shroyer, M.J., Bennett, S.E., Putnam, C.D., Tainer, J.A. and Mosbaugh, D.W. (1999) Mutation of an active site residue in Escherichia coli uracil-DNA glycosylase: effect on DNA binding, uracil inhibition and catalysis. *Biochemistry*, **38**, 4834-4845.

15. Barrett, T.E., Scharer, O.D., Savva, R., Brown, T., Jiricny, J., Verdine, G.L. and Pearl, L.H. (1999) Crystal structure of a thwarted mismatch glycosylase DNA repair complex. *EMBO J*, **18**, 6599-6609.

16. Kamiya, H. (2003) Mutagenic potentials of damaged nucleic acids produced by reactive oxygen/nitrogen species: approaches using synthetic oligonucleotides and nucleotides: survey and summary. *Nucleic Acids Res*, **31**, 517-531.

17. Case-Green, S.C. and Southern, E.M. (1994) Studies on the base pairing properties of deoxyinosine by solid phase hybridisation to oligonucleotides. *Nucleic Acids Res*, **22**, 131-136.

18. Martin, F.H., Castro, M.M., Aboul-ela, F. and Tinoco, I., Jr. (1985) Base pairing involving deoxyinosine: implications for probe design. *Nucleic Acids Res*, **13**, 8927-8938.

19. Watkins, N.E., Jr. and SantaLucia, J., Jr. (2005) Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes. *Nucleic Acids Res*, **33**, 6258-6267.

20. Parikh, S.S., Mol, C.D., Slupphaug, G., Bharati, S., Krokan, H.E. and Tainer, J.A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J*, **17**, 5214-5226.

21. Savva, R., McAuley-Hecht, K., Brown, T. and Pearl, L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487-493.

22. Drohat, A.C., Jagadeesh, J., Ferguson, E. and Stivers, J.T. (1999) Role of electrophilic and general base catalysis in the mechanism of Escherichia coli uracil DNA glycosylase. *Biochemistry*, **38**, 11866-11875.

23. Barrett, T.E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J. and Pearl, L.H. (1998) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell*, **92**, 117-129.

CHAPTER FIVE

RESEARCH SIGNIFICANCE AND CONCLUDING REMARKS

Cells are constantly attacked by endogenous and environment factors and thus it cause cell death or mutation by DNA damage. DNA repair mechanisms are required to remove DNA damage.

Among several factors that can cause damage in DNA are reactive nitrogen species (RNS). $NO^{\bullet}$ plays a role in the nitrosative stress as a main source of RNS produced by cells and/or as a main source for the other RNS such as nitrogen dioxide, dinitrogen trioxide, dinitrogen tetraoxide, nitrite, nitrosothiol and peroxynitrite (1-2). $NO^{\bullet}$ is also endogenously generated by NO synthases (NOS) from L-arginine and oxygen (3). To detoxify RNS, cells are able to decrease the damage using several proteins such as thioredoxins or thioredoxin domain containing proteins. In this study, we used a mouse lung cDNA library for finding candidate genes that involved in resistance to nitrosative stress using acidified nitrite as a source of RNS (4-5). Mouse thioredoxin domain-containing 5 (mTrx 5) was identified as one of the candidate genes. mTrx 5 complemented the thioredoxin deficiencies of the double deletion mutant in *E. coli* (Δ *trxA/*Δ*trxC*). Purified mTrx 5 proteins can also reduced DNA damage that is generated by RNS. This work takes advantage of a hypersensitive Escherichia coli genetic system to identify genes involved in resistance to nitrosative stress in mouse lungs.

Mismatch-specific uracil-DNA glycosylase (MUG) as a sequence homolog of the human thymine DNA glycosylase (TDG) that belongs to Family 2 in the UDG

superfamily that removes highly specific on G:U mismatches in double-stranded DNA (dsDNA) (6). Using cell extract from *E. coli* triple mutant strain (*nfi, nei, alkA*), we detected xanthine DNA glycosylase (XDG) activity. This work is the first report that *E. coli* MUG has robust XDG activity. Based on sequence alignment, we have chosen several amino acids to identify amino acid residues that may play a role in recognition of deaminated bases. It revealed that N140 and S23 are important determinants for XDG activity in *E. coli* MUG. Molecular modeling and molecular dynamics simulations indicate that distinct hydrogen bonding patterns in the active site of *E. coli* MUG, which account for the specificity differences between the wild type MUG and the N140 and S23 mutants.

Based on their similarity of sequence and structure, the uracil DNA glycosylase (UDG) superfamily consists of five families and having a Uracil repair activity regardless of different family. In this work, we found that a new family posseses no uracil DNA glycosylase activity from archara *Methanosarcina barkeri* (Mba). Surprisingly, the reombonant protein of Mba has strong activity on Hypoxanthine which is deamination product of adenine, and a low activity on xanthine, the deamination product of guanine. To verify the catalytic site, we chose several potential amino acids based human TDG guided-model of Mba protein. Among the several candidates, we identified only N39 is completely conserved in the Mba enzyme and its homologs. Notably when N39 was substituted with Ala, N39A completely abolished the glycosylase activity. Therefore N39 might play a key role as the catalytic residue in Mba enzyme. To further characterize the Mba enzyme, we used the lacZ-based genetic assay (7,8). *E. coli* CC106 strain contains

G to A transition mutation at codon 461 that allows for detection of A to G revertants. We cloned wt Mba and N39A mutant and transformed them into *nfi* deficient CC106 strain. While the N39A mutant failed to suppress A to G mutations in the CC106 *nfi*$^-$ cells under nitrosative stress, the wt Mba DNA glycosylase reduced the number of revertants substantially. These results indicate that the hypoxanthine DNA glycosylase activity (HDG) in the Mba repair enzyme is capable of removing hypoxanthine *in vivo*.

References

[1] Wink, D. A.; Mitchell, J. B. Chemical biology of nitric oxide: Insights into regulatory, cytotoxic, and cytoprotective mechanisms of nitric oxide. *Free Radic Biol Med* 25 (1998) 434-456.

[2] Nathan, C.; Shiloh, M. U. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc Natl Acad Sci U S A* 97 (2000) 8841-8848.

[3] Martínez MC, Andriantsitohaina R. Reactive nitrogen species: molecular mechanisms and potential significance in health and disease. *Antioxid Redox Signal*. 11 (2009) 669-702.

[4] Darwin, K. H.; Ehrt, S.; Gutierrez-Ramos, J. C.; Weich, N.; Nathan, C. F. The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. *Science* 302 (20003) 1963-1966.

[5] Venkatesh, J.; Kumar, P.; Krishna, P. S.; Manjunath, R.; Varshney, U. Importance of uracil DNA glycosylase in Pseudomonas aeruginosa and Mycobacterium smegmatis, G+C-rich bacteria, in mutation prevention, tolerance to acidified nitrite, and endurance in mouse macrophages. *J Biol Chem* 278 (2003) 24350-24358.

[6] P. Gallinari, J. Jiricny. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature* 383 (1996) 735–738.

[7] Cupples, C.G. and Miller, J.H. A set of lacZ mutations in Escherichia coli that allow rapid detection of each of the six base substitutions. *Proc Natl Acad Sci U S A* 86 (1989) 5345-5349.

[8] Josephy, P.D. The Escherichia coli lacZ reversion mutagenicity assay. *Mutat Res* (2000) 455 71-80.